

HP 5500 EI & 5500 SI Switch Series Configuration Guides

Software version: Release 2215
Document version: 6W100-20120428



Legal and notice information

© Copyright 2012 Hewlett-Packard Development Company, L.P.

No part of this documentation may be reproduced or transmitted in any form or by any means without prior written consent of Hewlett-Packard Development Company, L.P.

The information contained herein is subject to change without notice.

HEWLETT-PACKARD COMPANY MAKES NO WARRANTY OF ANY KIND WITH REGARD TO THIS MATERIAL, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. Hewlett-Packard shall not be liable for errors contained herein or for incidental or consequential damages in connection with the furnishing, performance, or use of this material.

The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

About the HP 5500 EI & 5500 SI configuration guides-Release 2215

The HP 5500 EI & 5500 SI configuration guides describe the software features for the HP 5500 EI & 5500 SI Switch Series, software release train 2215, and guide you through the software configuration procedures. These configuration guides also provide configuration examples to help you apply the software features to different network scenarios.

Configuration guide	Content
01 Fundamentals Configuration Guide	<p>Describes how to use the command line interface of the switch, log in to and set up the switch. This guide includes:</p> <ul style="list-style-type: none">• CLI (command line interface overview and how to use the CLI)• Login methods• CLI login• Web login• NMS login• User login control• FTP• TFTP• File system management• Configuration file management• Software upgrade• ISSU• ISSU configuration example• Device management• Automatic configuration
02 IRF Configuration Guide	<p>Describes the HP proprietary IRF technology, IRF fabric setup and configuration procedure, restrictions and guidelines, and IRF fabric configuration procedure.</p>
03 Layer 2 – LAN Switching Configuration Guide	<p>Covers Layer 2 technologies and features used on a LAN switched network, including VLAN technology, port isolation, and spanning tree. You can use these features to divide broadcast domains, remove Layer 2 loops, isolate users within a VLAN, re-mark VLAN tags, implement VLAN VPNs over the Internet, and so on. This guide includes:</p> <ul style="list-style-type: none">• Ethernet interface• Loopback and null interfaces• Bulking configuring interfaces• MAC address table• MAC Information• Ethernet link aggregation• Port isolation• Spanning tree• BPDU tunneling• VLAN

Configuration guide	Content
	<ul style="list-style-type: none"> • Super VLAN • Isolate-user-VLAN • Voice VLAN • GVRP • QinQ • VLAN mapping • LLDP • Service loopback group • MVRP
04 Layer 3 – IP Services Configuration Guide	<p>Describes IP addressing (including static and dynamic IPv4 and IPv6 address assignment), IRDP, UDP helper, DNS, network performance optimization, ARP, and interoperation between IPv4 and IPv6. This guide includes:</p> <ul style="list-style-type: none"> • ARP • Gratuitous ARP • Proxy ARP • ARP snooping • IP addressing • DHCP overview • DHCP server • DHCP relay agent • DHCP client • DHCP snooping • BOOTP client • IPv4 DNS • IRDP • IP performance optimization • UDP helper • IPv6 basics • DHCPv6 overview • DHCPv6 server • DHCPv6 relay agent • DHCPv6 client • DHCPv6 snooping • IPv6 DNS • Tunneling
05 Layer 3 – IP Routing Configuration Guide	<p>Covers the routing technologies for IPv4 and IPv6 networks of different sizes, route filtering, route control, and policy based routing. This guide includes:</p> <ul style="list-style-type: none"> • IP routing basics • Static routing • RIP • OSPF • IS-IS • BGP • IPv6 static routing • RIPng • OSPFv3 • IPv6 IS-IS • IPv6 BGP

Configuration guide	Content
	<ul style="list-style-type: none"> • Routing policy • Policy-based routing • MCE • IPv6 MCE
06 IP Multicast Configuration Guide	<p>Covers Layer 3 IPv4 multicast protocols (IGMP, PIM, MSDP, and MBGP), Layer 3 IPv6 multicast protocols (MLD, IPv6 PIM, and IPv6 MBGP), Layer 2 IPv4 multicast protocols (IGMP snooping, PIM snooping, and multicast VLAN), and Layer 2 IPv6 multicast protocols (MLD snooping, IPv6 PIM snooping, and IPv6 multicast VLAN). This guide includes:</p> <ul style="list-style-type: none"> • Multicast overview • IGMP snooping • PIM snooping • Multicast VLAN • Multicast routing and forwarding • IGMP • PIM • MSDP • MBGP • MLD snooping • IPv6 PIM snooping • IPv6 multicast VLAN • IPv6 multicast routing and forwarding • MLD • IPv6 PIM • IPv6 MBGP
07 ACL and QoS Configuration Guide	<p>Describes how to classify traffic with ACLs, and allocate network resources and manage congestions with QoS technologies to improve network performance and network use efficiency. You can use ACLs to help other function modules (such as QoS and IP routing) classify or filter traffic. This guide includes:</p> <ul style="list-style-type: none"> • ACL • QoS overview • QoS configuration approaches • Priority mapping • Traffic policing, traffic shaping, and line rate • Congestion management • Congestion avoidance • Traffic filtering • Priority marking • Traffic redirecting • Class-based accounting • Data buffer • Appendix

Configuration guide	Content
08 Security Configuration Guide	<p>Covers security features. The major security features available on the switch include identity authentication (AAA), access security (802.1X, MAC authentication, portal, and port security), secure management (SSH), and attack protection (IP source guard, ARP attack protection, and URPF). This guide includes:</p> <ul style="list-style-type: none">• AAA• 802.1X fundamentals• 802.1X• EAD fast deployment• MAC authentication• Portal• Triple authentication• Port security• User profile• Password control• HABP• Public key• PKI• IPsec• SSH2.0• SFTP• SCP• SSL• TCP attack protection• IP source guard• ARP attack protection• ND attack defense• URPF• SAVI• Blacklist
09 High Availability Configuration Guide	<p>Describes high availability technologies and features available on the switch for failure detection and failover. Failure detection technologies focus on fault detection and isolation. Failover technologies focus on network recovery. This guide includes:</p> <ul style="list-style-type: none">• High availability overview• Ethernet OAM• CFD• DLDP• RRPP• Smart Link• Monitor Link• VRRP• Stateful failover• BFD• Track

Configuration guide	Content
10 Network Management and Monitoring Configuration Guide	<p data-bbox="608 230 1422 383">Describes features that help you manage and monitor your network, for example, manage system events, sample packets, assess network performance, synchronize the clock for all devices with the clock in the network, supply power for attached devices by using PoE, and test network connectivity. This guide includes:</p> <ul data-bbox="608 398 1023 887" style="list-style-type: none"> <li data-bbox="608 398 1023 432">• System maintenance and debugging <li data-bbox="608 439 687 472">• NTP <li data-bbox="608 479 831 512">• Information center <li data-bbox="608 519 711 553">• SNMP <li data-bbox="608 560 719 593">• RMON <li data-bbox="608 600 783 633">• Port mirroring <li data-bbox="608 640 807 674">• Traffic mirroring <li data-bbox="608 680 703 714">• NQA <li data-bbox="608 721 703 754">• sFlow <li data-bbox="608 761 679 795">• IPC <li data-bbox="608 801 679 835">• PoE <li data-bbox="608 842 935 875">• Cluster managementCWMP <li data-bbox="608 882 695 916">• Stack

Contents

Using the CLI	1
Logging in to the CLI	1
Command conventions	1
Using the undo form of a command	2
CLI views	2
Entering system view from user view	3
Returning to the upper-level view from any view	3
Returning to user view from any other view	4
Accessing the CLI online help	4
Entering a command	5
Editing a command line	5
Entering a STRING type value for an argument	5
Abbreviating commands	6
Configuring and using command keyword aliases	6
Configuring and using hotkeys	6
Enabling redisplaying entered-but-not-submitted commands	8
Understanding command-line error messages	8
Using the command history function	9
Viewing history commands	9
Setting the command history buffer size for user interfaces	9
Controlling the CLI output	10
Pausing between screens of output	10
Filtering the output from a display command	10
Configuring user privilege and command levels	13
Configuring a user privilege level	13
Switching the user privilege level	17
Changing the level of a command	19
Saving the running configuration	19
Displaying and maintaining CLI	19
Login overview	21
Login methods at a glance	21
User interfaces	22
User interface assignment	22
User interface numbering	22
Logging in to the CLI	23
Logging in through the console port for the first time	23
Configuring console login control settings	25
Configuring none authentication for console login	26
Configuring password authentication for console login	27
Configuring scheme authentication for console login	28
Configuring common console login settings (optional)	30
Logging in through Telnet	32
Configuring none authentication for Telnet login	33
Configuring password authentication for Telnet login	34
Configuring scheme authentication for Telnet login	35
Configuring common settings for VTY user interfaces (optional)	37
Using the device to log in to a Telnet server	39
Setting the DSCP value for IP to use for outgoing Telnet packets	40

Logging in through SSH	40
Configuring the SSH server on the device	40
Using the device as an SSH client to log in to the SSH server	43
Modem dial-in through the console port	43
Setting up the configuration environment	44
Configuring none authentication for modem dial-in	47
Configuring password authentication for modem dial-in	47
Configuring scheme authentication for modem dial-in	48
Configuring common settings for modem dial-in (optional)	51
Displaying and maintaining CLI login	53
Logging in to the Web interface	54
Configuring HTTP login	54
Configuring HTTPS login	55
Displaying and maintaining Web login	57
HTTP login configuration example	57
Network requirements	57
Configuration procedure	57
HTTPS login configuration example	59
Network requirements	59
Configuration procedure	59
Logging in through NMS	61
Configuring SNMP login	61
Prerequisites	61
Configuring SNMPv3 settings	61
Configuring SNMPv1 or SNMPv2c settings	62
NMS login example	63
Network requirements	63
Configuration procedure	63
Controlling user logins	65
Controlling Telnet logins	65
Configuring source IP-based Telnet login control	65
Configuring source/destination IP-based Telnet login control	66
Configuring source MAC-based Telnet login control	66
Telnet login control configuration example	66
Configuring source IP-based SNMP login control	67
Configuration procedure	67
SNMP login control configuration example	68
Configuring Web login control	69
Configuring source IP-based Web login control	69
Logging off online Web users	69
Web login control configuration example	69
Configuring FTP	71
Using the device as an FTP client	71
Establishing an FTP connection	71
Setting the DSCP value for IP to use for outgoing FTP packets	72
Managing directories on the FTP server	73
Working with the files on the FTP server	73
Switching to another user account	74
Maintaining and troubleshooting the FTP connection	74
Terminating the FTP connection	74
FTP client configuration example	75
Using the device as an FTP server	76

Configuring basic parameters	76
Configuring authentication and authorization	77
FTP server configuration example	78
Displaying and maintaining FTP	80
Configuring TFTP	81
Prerequisites	81
Using the device as a TFTP client	81
Displaying and maintaining the TFTP client	82
TFTP client configuration example	82
Managing the file system	84
Managing files	84
Displaying file information	85
Displaying file contents	85
Renaming a file	85
Copying a file	85
Moving a file	86
Deleting/restoring a file	86
Emptying the recycle bin	86
Managing directories	86
Displaying directory information	86
Displaying the current working directory	87
Changing the current working directory	87
Creating a directory	87
Removing a directory	87
Managing storage media	87
Managing storage medium space	87
Performing batch operations	88
Setting the file system operation mode	88
File system management examples	88
Managing configuration files	90
Overview	90
Configuration types	90
Configuration file format and content	91
Coexistence of multiple configuration files	91
Startup with a configuration file	91
Saving the running configuration	91
Enabling configuration file auto-update	92
Saving running configuration in fast mode or safe mode	92
Configuring configuration rollback	93
Configuration task list	93
Configuring configuration archive parameters	93
Enabling automatic configuration archiving	95
Manually archiving running configuration	95
Performing configuration rollback	95
Specifying a configuration file for the next startup	96
Backing up the next-startup configuration file to a TFTP server	96
Deleting the next-startup configuration file	97
Restoring the next-startup configuration file from a TFTP server	97
Displaying and maintaining a configuration file	98
Upgrading software	99
Software upgrade methods	99
Upgrading software through a system reboot	100

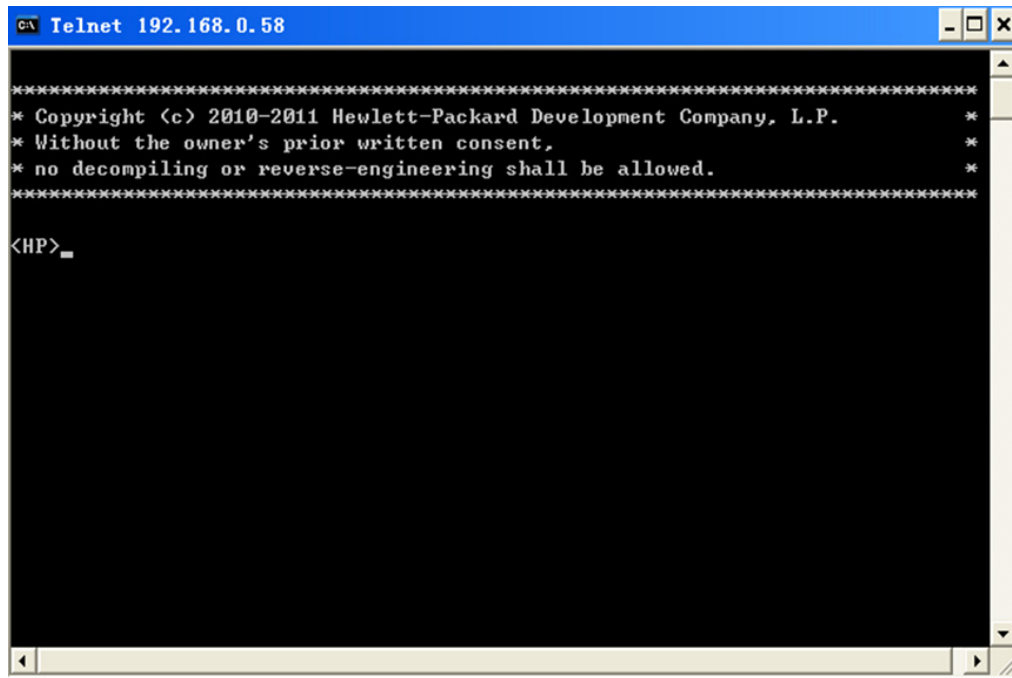
Upgrading Boot ROM through a system reboot	100
Upgrading system software through system reboot (method 1)	100
Upgrading system software through system reboot (method 2)	101
Upgrading software by installing hotfixes	101
Basic concepts	101
Patch state	102
Hotfix configuration task list	104
Installation prerequisites	105
Installing a patch in one step	105
Installing a patch step-by-step	106
Uninstalling a patch step-by-step	107
Displaying and maintaining software upgrade	108
Software upgrade examples	108
Immediate upgrade configuration example	108
Hotfix configuration example	110
Performing ISSU	112
ISSU overview	112
ISSU upgrade procedure	112
ISSU states	113
System software version rollback	114
Performing an ISSU	114
ISSU upgrade task list	114
ISSU upgrade prerequisites	115
Displaying version compatibility	116
Performing an ISSU for a compatible version	116
Performing an ISSU for an incompatible version	117
Setting the ISSU version rollback timer	118
Performing a manual version rollback	118
Displaying and maintaining ISSU	119
ISSU upgrade example	120
Network status	120
Network requirements	120
Upgrade procedure	121
Configuring link aggregation	121
Configuring the TFTP server	123
Downloading the new system software image	123
Checking all IRF member switches before the ISSU upgrade	124
Viewing the version compatibility	126
Performing compatible ISSU upgrade	127
Performing incompatible ISSU upgrade	128
Managing the device	130
Configuring the device name	130
Changing the system time	130
Configuration guidelines	130
Configuration procedure	133
Enabling displaying the copyright statement	133
Changing the brand name	134
Configuration preparation	134
Configuration guidelines	135
Configuration procedure	135
Configuring banners	135
Banner message input modes	136
Configuration procedure	137

Configuring the exception handling method.....	137
Rebooting the device	137
Rebooting devices immediately at the CLI	138
Scheduling a device reboot	138
Scheduling jobs.....	139
Job configuration approaches	139
Configuration guidelines	139
Scheduling a job in the non-modular approach	140
Scheduling a job in the modular approach	140
Disabling Boot ROM access	141
Configuring the port status detection timer.....	141
Configuring temperature thresholds for a device	141
Clearing unused 16-bit interface indexes.....	142
Verifying and diagnosing transceiver modules	142
Diagnosing transceiver modules.....	143
Displaying and maintaining device management.....	143
Automatic configuration introduction.....	145
Typical application scenario.....	145
How automatic configuration works	146
Automatic configuration work flow	146
Using DHCP to obtain an IP address and other configuration information	147
Obtaining the configuration file from the TFTP server.....	148
Executing the configuration file.....	150
Index	151

Using the CLI

At the command-line interface (CLI), you can enter text commands to configure, manage, and monitor your device.

Figure 1 CLI example



Logging in to the CLI

You can log in to the CLI in a variety of ways. For example, you can log in through the console port, or by using Telnet or SSH. For more information about login methods, see "Logging in to the CLI "

Command conventions

Command conventions help you understand the syntax of commands. Commands in product manuals comply with the conventions listed in [Table 1](#).

Table 1 Command conventions

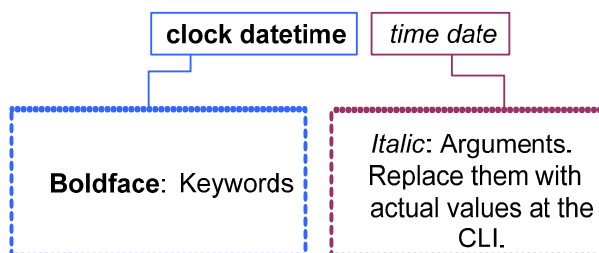
Convention	Description
Boldface	Bold text represents commands and keywords that you enter literally as shown.
<i>Italic</i>	<i>Italic</i> text represents arguments that you replace with actual values.
[]	Square brackets enclose syntax choices (keywords or arguments) that are optional.
{ x y ... }	Braces enclose a set of required syntax choices separated by vertical bars, from which you select one.

Convention	Description
[x y ...]	Square brackets enclose a set of optional syntax choices separated by vertical bars, from which you select one or none.
{ x y ... } *	Asterisk marked braces enclose a set of required syntax choices separated by vertical bars, from which you select at least one.
[x y ...] *	Asterisk marked square brackets enclose optional syntax choices separated by vertical bars, from which you select one choice, multiple choices, or none.
&<1-n>	The argument or keyword and argument combination before the ampersand (&) sign can be entered 1 to n times.
#	A line that starts with a pound (#) sign is comments.

Command keywords are case insensitive.

The following example analyzes the syntax of the **clock datetime** *time date* command according to [Table 1](#).

Figure 2 Understanding command-line parameters



For example, to set the system time to 10:30:20, February 23, 2011, enter the following command line at the CLI and press **Enter**:

```
<Sysname> clock datetime 10:30:20 2/23/2011
```

Using the undo form of a command

Most configuration commands have an **undo** form for canceling a configuration, restoring the default, or disabling a feature. For example, the **info-center enable** command enables the information center, and the **undo info-center enable** command disables the information center.

CLI views

Commands are grouped in different views by function. To use a command, you must enter the view of the command.

CLI views are organized in a hierarchical structure, as shown in [Figure 3](#). Each view has a unique prompt, from which you can identify where you are and what you can do. For example, the prompt [Sysname-vlan100] shows that you are in the view of VLAN 100 and can configure attributes for the VLAN.

You are placed in user view immediately after you are logged in to the CLI. The user view prompt is <Device-name>, where the *Device-name* argument defaults to **HP** and can be changed by using the **sysname** command. In user view, you can perform some basic operations, including display, debug, file

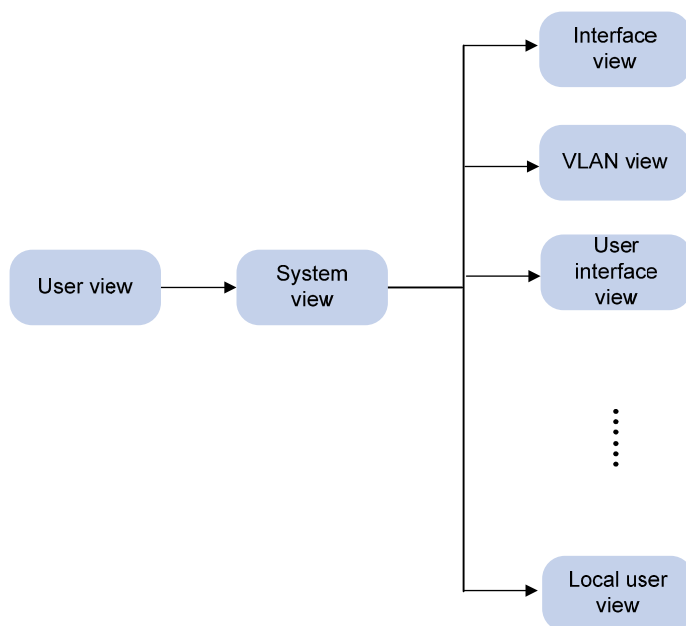
management, FTP, Telnet, clock setting, and reboot. For more information about the **sysname** command, see *Fundamentals Command Reference*.

From user view, you can enter system view to configure global settings, including the daylight saving time, banners, and hotkeys. The system view prompt is `[Device-name]`.

From system view, you can enter different function views. For example, you can enter interface view to configure interface parameters, enter VLAN view to add ports to the specific VLAN, enter user interface view to configure login user attributes, or create a local user and enter local user view to configure attributes for the local user.

To display all commands available in a view, enter a question mark (?) at the view prompt.

Figure 3 CLI view hierarchy



Entering system view from user view

Task	Command
Enter system view from user view.	system-view

Returning to the upper-level view from any view

Task	Command
Return to the upper-level view from any view.	quit

Executing the **quit** command in user view terminates your connection to the device.

NOTE:

In public key code view, use the **public-key-code end** command to return to the upper-level view (public key view). In public key view, use the **peer-public-key end** command to return to system view.

Returning to user view from any other view

You can return to user view from any other view by using the **return** command, instead of using the **quit** command repeatedly. Pressing **Ctrl+Z** has the same effect.

To return to user view from any other view:

Task	Command
Return to user view.	return

Accessing the CLI online help

The CLI online help is context sensitive. You can enter a question mark at any point of a command to display all available options.

To access the CLI online help, use one of the following methods:

- Enter a question mark at a view prompt to display the first keywords of all commands available in the view. For example:

```
<Sysname> ?
User view commands:
  archive          Specify archive settings
  backup           Backup next startup-configuration file to TFTP server
  boot-loader      Set boot loader
  bootrom          Update/read/backup/restore bootrom
  cd               Change current directory
...

```

- Enter some keywords of a command and a question mark separated by a space to display available keywords and arguments.
 - Example 1: The question mark is in the place of a keyword, and the CLI displays all possible keywords with a brief description for each keyword.

```
<Sysname> terminal ?
  debugging  Send debug information to terminal
  logging    Send log information to terminal
  monitor    Send information output to current terminal
  trapping   Send trap information to terminal

```

- Example 2: The question mark is in the place of an argument, and the CLI displays the description of the argument.

```
<Sysname> system-view
[Sysname] interface vlan-interface ?
  <1-4094>  VLAN interface
[Sysname] interface vlan-interface 1 ?
  <cr>
[Sysname] interface vlan-interface 1

```

The string **<cr>** indicates that the command is complete, and you can press **Enter** to execute the command.

- Enter an incomplete keyword string followed by a question mark to display all keywords starting with the string. For example:


```

<Sysname> f?
    fixdisk
    format
    free
    ftp
<Sysname> display ftp?
    ftp
    ftp-server
    ftp-user

```

Entering a command

When you enter a command, you can use some keys or hotkeys to edit the command line, or use abbreviated keywords or keyword aliases.

Editing a command line

You can use the keys listed in [Table 2](#) or the hotkeys listed in [Table 3](#) to edit a command line.

Table 2 Keys for editing a command line

Key	Function
Common keys	If the edit buffer is not full, pressing a common key inserts the character at the position of the cursor and moves the cursor to the right.
Backspace	Deletes the character to the left of the cursor and moves the cursor back one character.
Left arrow key or Ctrl+B	Moves the cursor one character to the left.
Right arrow key or Ctrl+F	Moves the cursor one character to the right.
Tab	<p>If you press Tab after entering part of a keyword, the system automatically completes the keyword:</p> <ul style="list-style-type: none"> • If a unique match is found, the system substitutes the complete keyword for the incomplete one and displays what you entered in the next line. • If there is more than one match, you can press Tab repeatedly to choose the keyword you want to enter. • If there is no match, the system does not modify what you entered but displays it again in the next line.

Entering a STRING type value for an argument

Generally, a STRING type argument value can contain any printable character (in the ASCII code range of 32 to 126) other than the question mark (?), quotation mark ("), backward slash (\), and space. However, a specific STRING type argument might have more strict requirements. For example, the domain name is of the STRING type. Invalid characters for it include the vertical bar (|), slash (/), colon (:), asterisk (*), less-than sign (<), greater-than sign (>), and at sign (@), as well as the question mark (?), quotation mark ("), backward slash (\), and space. For more information about the specific requirements for a STRING type argument, see the relevant command reference.

```

<Sysname> system-view
[Sysname] domain ?

```

STRING<1-24> Domain name

Abbreviating commands

You can enter a command line quickly by entering incomplete keywords that can uniquely identify the complete command. In user view, for example, commands starting with an **s** include **startup saved-configuration** and **system-view**. To enter system view, you only need to enter **sy**. To set the configuration file to be used at the next startup, you can enter **st s**.

You can also press **Tab** to have an incomplete keyword automatically completed.

Configuring and using command keyword aliases

The command keyword alias function allows you to replace the first keyword of a non-undo command or the second keyword of an **undo** command with your preferred keyword when you execute the command. For example, if you configure **show** as the alias for the **display** keyword, you can enter **show** to execute a **display** command.

Usage guidelines

- After you successfully execute a command by using a keyword alias, the system saves the keyword, instead of its alias, to the running configuration.
- If you press **Tab** after entering part of an alias, the keyword is displayed.
- If a string you entered partially matches a keyword and an alias, the command indicated by the alias is executed. To execute the command indicated by the keyword, enter the complete keyword.
- If a string you entered exactly matches a keyword but partially matches an alias, the command indicated by the keyword is executed. To execute the command indicated by the alias, enter the complete alias.
- If you enter a string that partially matches multiple aliases, the system gives you a prompt.

Configuration procedure

To configure a command keyword alias:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable the command keyword alias function.	command-alias enable	By default, the command keyword alias function is disabled.
3. Configure a command keyword alias.	command-alias mapping <i>cmdkey</i> <i>alias</i>	By default, no command keyword alias is configured. You must enter the <i>cmdkey</i> and <i>alias</i> arguments in their complete form.

Configuring and using hotkeys

To facilitate CLI operation, the system defines some hotkeys and provides five configurable command hotkeys. Pressing a command hotkey equals entering a command. For system-reserved hotkeys, see [Table 3](#).

To configure hotkeys:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure hotkeys.	hotkey { CTRL_G CTRL_L CTRL_O CTRL_T CTRL_U } <i>command</i>	By default: <ul style="list-style-type: none"> • Ctrl+G is assigned the display current-configuration command. • Ctrl+L is assigned the display ip routing-table command. • Ctrl+O is assigned the undo debugging all command. No command is assigned to Ctrl+T or Ctrl+U .
3. Display hotkeys.	display hotkey [{ begin exclude include } <i>regular-expression</i>]	Optional. Available in any view. See Table 3 for hotkeys reserved by the system.

The hotkeys in Table 3 are defined by the device. If a hotkey is also defined by the terminal software that you are using to interact with the device, the definition of the terminal software takes effect.

Table 3 Hotkeys reserved by the system

Hotkey	Function
Ctrl+A	Moves the cursor to the beginning of the line.
Ctrl+B	Moves the cursor one character to the left.
Ctrl+C	Stops the current command.
Ctrl+D	Deletes the character at the cursor.
Ctrl+E	Moves the cursor to the end of the line.
Ctrl+F	Moves the cursor one character to the right.
Ctrl+H	Deletes the character to the left of the cursor.
Ctrl+K	Aborts the connection request.
Ctrl+N	Displays the next command in the command history buffer.
Ctrl+P	Displays the previous command in the command history buffer.
Ctrl+R	Redisplays the current line.
Ctrl+V	Pastes text from the clipboard.
Ctrl+W	Deletes the word to the left of the cursor.
Ctrl+X	Deletes all characters to the left of the cursor.
Ctrl+Y	Deletes all characters to the right of the cursor.
Ctrl+Z	Returns to user view.
Ctrl+]	Terminates an incoming connection or a redirect connection.
Esc+B	Moves the cursor back one word.
Esc+D	Deletes all characters from the cursor to the end of the word.
Esc+F	Moves the cursor forward one word.

Hotkey	Function
Esc+N	Moves the cursor down one line (available before you press Enter)
Esc+P	Moves the cursor up one line (available before you press Enter)
Esc+<	Moves the cursor to the beginning of the clipboard.
Esc+>	Moves the cursor to the ending of the clipboard.

Enabling redisplaying entered-but-not-submitted commands

After you enable redisplaying entered-but-not-submitted commands:

- If you entered nothing at the command-line prompt before the system outputs system information such as logs, the system does not display the command-line prompt after the output.
- If you entered some information (except Yes or No for confirmation), the system displays a line break and then display what you have entered after the output.

To enable redisplaying entered-but-not-submitted commands:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable redisplaying entered-but-not-submitted commands.	info-center synchronous	By default, the feature is disabled. For more information about this command, see <i>Network Management and Monitoring Command Reference</i> .

Understanding command-line error messages

If a command line fails the syntax check, the CLI displays error messages.

Table 4 Common command-line error messages

Error message	Cause
% Unrecognized command found at '^' position.	The keyword in the marked position is invalid.
% Incomplete command found at '^' position.	One or more required keywords or arguments are missing.
% Ambiguous command found at '^' position.	The entered character sequence matches more than one command.
Too many parameters	The entered character sequence contains excessive keywords or arguments.
% Wrong parameter found at '^' position.	The argument in the marked position is invalid.

Using the command history function

The system can automatically save successfully executed commands to the command history buffer for the current user interface. You can view them and execute them again, or set the maximum number of commands that can be saved in the command history buffer.

A command is saved to the command history buffer in the exact format as it was entered. For example, if you enter an incomplete command, the command saved in the command history buffer is also incomplete; if you enter a command by using a command keyword alias, the command saved in the command history buffer also uses the alias.

If you enter a command in the same format repeatedly in succession, the system buffers the command only once. If you enter a command repeatedly in different formats, the system buffers each command format. For example, **display cu** and **display current-configuration** are buffered as two entries but successive repetitions of **display cu** create only one entry in the buffer.

By default, the command history buffer can save up to 10 commands for each user. To set the capacity of the command history buffer for the current user interface, use the **history-command max-size** command.

Viewing history commands

You can use arrow keys to access history commands in Windows 200x and Windows XP Terminal or Telnet. In Windows 9x HyperTerminal, the arrow keys are invalid, and you must use **Ctrl+P** and **Ctrl+N** instead.

To view command history, use one of the following methods:

Task	Command
Display all commands in the command history buffer.	display history-command [{ begin exclude include } <i>regular-expression</i>]
Display the previous history command.	Up arrow key or Ctrl+P
Display the next history command.	Down arrow key or Ctrl+N

Setting the command history buffer size for user interfaces

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter user interface view.	user-interface { <i>first-num1</i> [<i>last-num1</i>] { aux vty } <i>first-num2</i> [<i>last-num2</i>] }	N/A
3. Set the maximum number of commands that can be saved in the command history buffer.	history-command max-size <i>size-value</i>	Optional. By default, the command history buffer can save up to 10 commands.

Controlling the CLI output

This section describes the CLI output control features that help you quickly identify the desired output.

Pausing between screens of output

If the output being displayed is more than will fit on one screen, the system automatically pauses after displaying a screen. By default, up to 24 lines can be displayed on a screen. To change the screen length, use the **screen-length** *screen-length* command. For more information about this command, see *Fundamentals Command Reference*. To control output, use keys in Table 5.

Table 5 Keys for controlling output

Keys	Function
Space	Displays the next screen.
Enter	Displays the next line.
Ctrl+C	Stops the display and cancels the command execution.
<PageUp>	Displays the previous page.
<PageDown>	Displays the next page.

To display all output at one time and refresh the screen continuously until the last screen is displayed:

Task	Command	Remarks
Disable pausing between screens of output for the current session.	screen-length disable	<p>The default for a session depends on the setting of the screen-length command in user interface view. The default of the screen-length command is pausing between screens of output and displaying up to 24 lines on a screen.</p> <p>This command is executed in user view, and takes effect only for the current session. When you relog in to the device, the default is restored.</p>

Filtering the output from a display command

You can use one of the following methods to filter the output from a **display** command:

- Specify the | { **begin** | **exclude** | **include** } *regular-expression* option at the end of the command.
- When the system pauses after displaying a screen of output, enter a forward slash (/), minus sign (-), or plus sign (+) plus a regular expression to filter subsequent output. The forward slash equals the keyword **begin**, the minus sign equals the keyword **exclude**, and the plus sign equals the keyword **include**.

The following definitions apply to the **begin**, **exclude**, and **include** keywords:

- **begin**—Displays the first line that matches the specified regular expression and all lines that follow.
- **exclude**—Displays all lines that do not match the specified regular expression.
- **include**—Displays all lines that match the specified regular expression.

A regular expression is a case-sensitive string of 1 to 256 characters that supports the special characters in Table 6.

Table 6 Special characters supported in a regular expression

Character	Meaning	Remarks
^string	Starting sign. Matches a line that starts with <i>string</i> .	For example, regular expression "^user" matches a line beginning with "user", not "Auser".
string\$	Ending sign. Matches a line that ends with <i>string</i> .	For example, regular expression "user\$" only matches a line ending with "user", not "userA".
.	Matches any single character, such as a single character, a special character, and a blank.	For example, ".s" matches both "as" and "bs".
*	Matches the preceding character or character group zero or multiple times.	For example, "zo*" matches "z" and "zoo"; "(zo)*" matches "zo" and "zozo".
+	Matches the preceding character or character group one or multiple times	For example, "zo+" matches "zo" and "zoo", but not "z".
	Matches the preceding or succeeding character string	For example, "def int" only matches a character string containing "def" or "int".
_	If it is at the beginning or the end of a regular expression, it equals ^ or \$. In other cases, it equals comma, space, round bracket, or curly bracket.	For example, "a_b" matches "a b" or "a(b"; "_ab" only matches a line starting with "ab"; "ab_" only matches a line ending with "ab".
-	It connects two values (the smaller one before it and the bigger one after it) to indicate a range together with [].	For example, "1-9" means 1 to 9 (inclusive); "a-h" means a to h (inclusive).
[]	Matches a single character contained within the brackets.	For example, [16A] matches a string containing any character among 1, 6, and A; [1-36A] matches a string containing any character among 1, 2, 3, 6, and A (- is a hyphen). "]" can be matched as a common character only when it is put at the beginning of characters within the brackets, for example []string]. There is no such limit on "[".
()	A character group. It is usually used with "+" or "*".	For example, (123A) means a character group "123A"; "408(12)+" matches 40812 or 408121212. But it does not match 408.
\index	Repeats the character string specified by the index. A character string refers to the string within () before \. <i>index</i> refers to the sequence number (starting from 1 from left to right) of the character group before \. If only one character group appears before \, <i>index</i> can only be 1; if n character groups appear before <i>index</i> , <i>index</i> can be any integer from 1 to n.	For example, (string)\1 repeats <i>string</i> , and a matching string must contain <i>stringstring</i> . (string1)(string2)\2 repeats <i>string2</i> , and a matching string must contain <i>string1string2string2</i> . (string1)(string2)\1\2 repeats <i>string1</i> and <i>string2</i> respectively, and a matching string must contain <i>string1string2string1string2</i> .

Character	Meaning	Remarks
[^]	Matches a single character not contained within the brackets.	For example, [^16A] means to match a string containing any character except 1, 6 or A, and the matching string can also contain 1, 6 or A, but cannot contain only these three characters. For example, [^16A] matches "abc" and "m16", but not 1, 16, or 16A.
\<string	Matches a character string starting with <i>string</i> .	For example, "\<do" matches word "domain" and string "doa".
string\>	Matches a character string ending with <i>string</i> .	For example, "do\>" matches word "undo" and string "abcdo".
\bcharacter2	Matches <i>character1character2</i> . <i>character1</i> can be any character except number, letter or underline, and \b equals [^A-Za-z0-9_].	For example, "\ba" matches "a" with "-" being <i>character1</i> , and "a" being <i>character2</i> , but it does not match "2a" or "ba".
\Bcharacter	Matches a string containing <i>character</i> , and no space is allowed before <i>character</i> .	For example, "\Bt" matches "t" in "install", but not "t" in "big top".
character1\w	Matches <i>character1character2</i> . <i>character2</i> must be a number, letter, or underline, and \w equals [A-Za-z0-9_].	For example, "v\w" matches "vlan" ("v" is <i>character1</i> and "l" is <i>character2</i>) and "service" ("i" is <i>character2</i>).
\W	Equals \b.	For example, "\Wa" matches "a", with "-" being <i>character1</i> , and "a" being <i>character2</i> , but does not match "2a" or "ba".
\	Escape character. If a special character listed in this table follows \, the specific meaning of the character is removed.	For example, "\\\" matches a string containing "\", "\\^" matches a string containing "^", and "\\b" matches a string containing "b".

The following are several regular expression examples:

Use | **begin user-interface** in the **display current-configuration** command to match the first line of output that contains **user-interface** to the last line of output.

```
<Sysname> display current-configuration | begin user-interface
user-interface aux 0
user-interface vty 0 15
  authentication-mode none
  user privilege level 3
#
return
```

Use | **exclude Direct** in the **display ip routing-table** command to filter out direct routes and display only the non-direct routes.

```
<Sysname> display ip routing-table | exclude Direct
Routing Tables: Public
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
1.1.1.0/24	Static	60	0	192.168.0.0	Vlan1

Use | **include Vlan** in the **display ip routing-table** command to filter in route entries that contain **Vlan**.

```
<Sysname> display ip routing-table | include Vlan
```

```
Routing Tables: Public
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
192.168.1.0/24	Direct	0	0	192.168.1.42	Vlan999

Configuring user privilege and command levels

To avoid unauthorized access, the device defines the user privilege levels and command levels in Table 7. User privilege levels correspond to command levels. A user who has been logged in with a specific privilege level can use only the commands at that level or lower levels.

All commands are categorized into four levels: visit, monitor, system, and manage, and are identified from low to high, respectively by 0 through 3.

Table 7 Command levels and user privilege levels

Level	Privilege	Default set of commands
0	Visit	Includes commands for network diagnosis and commands for accessing an external device. Configuration of commands at this level cannot survive a device restart. Upon device restart, the commands at this level are restored to the default settings. Commands at this level include ping , tracert , telnet and ssh2 .
1	Monitor	Includes commands for system maintenance and service fault diagnosis. Commands at this level are not saved after being configured. After the device is restarted, the commands at this level are restored to the default settings. Commands at this level include debugging , terminal , refresh , and send .
2	System	Includes service configuration commands, including routing configuration commands and commands for configuring services at different network levels. By default, commands at this level include all configuration commands except for those at manage level.
3	Manage	Includes commands that influence the basic operation of the system and commands for configuring system support modules. By default, commands at this level involve the configuration commands of file system, FTP, TFTP, Xmodem download, user management, level setting, and parameter settings within a system (which are not defined by any protocols or RFCs).

Configuring a user privilege level

If the authentication mode on a user interface is scheme, configure a user privilege level for users who access the interface by using the AAA module or directly on the user interface. For SSH users who use public-key authentication, the user privilege level configured directly on the user interface always takes effect. For other users, the user privilege level configured in the AAA module has priority over the one configured directly on the user interface.

If the authentication mode on a user interface is none or password, configure the user privilege level directly on the user interface.

For more information about user login authentication, see "Logging in to the CLI." For more information about AAA and SSH, see *Security Configuration Guide*.

Configuring a user privilege level for users by using the AAA module

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter user interface view.	user-interface { <i>first-num1</i> [<i>last-num1</i>] { aux vty } <i>first-num2</i> [<i>last-num2</i>] }	N/A
3. Specify the scheme authentication mode.	authentication-mode scheme	By default, the authentication mode for VTY users is password , and no authentication is needed for AUX users.
4. Return to system view.	quit	N/A
5. Configure the authentication mode for SSH users as password .	For more information, see <i>Security Configuration Guide</i> .	This task is required only for SSH users who are required to provide their usernames and passwords for authentication.
6. Configure the user privilege level by using the AAA module.	<ul style="list-style-type: none"> • To use local authentication: <ol style="list-style-type: none"> a. Use the local-user command to create a local user and enter local user view. b. Use the level keyword in the authorization-attribute command to configure the user privilege level. • To use remote authentication (RADIUS or HWTACACS): Configure the user privilege level on the authentication server 	<p>User either approach.</p> <p>For local authentication, if you do not configure the user privilege level, the user privilege level is 0.</p> <p>For remote authentication, if you do not configure the user privilege level, the user privilege level depends on the default configuration of the authentication server.</p> <p>For more information about the local-user and authorization-attribute commands, see <i>Security Command Reference</i>.</p>

For example:

Configure the device to use local authentication for Telnet users on VTY 1 and set the user privilege level to 3.

```
<Sysname> system-view
[Sysname] user-interface vty 1
[Sysname-ui-vty1] authentication-mode scheme
[Sysname-ui-vty1] quit
[Sysname] local-user test
[Sysname-luser-test] password simple 123
[Sysname-luser-test] service-type telnet
```

When users Telnet to the device through VTY 1, they must enter username **test** and password **12345678**. After passing the authentication, the users can only use level-0 commands of level 0.

Assign commands of levels 0 through 3 to the users.

[Sysname-user-test] authorization-attribute level 3

Configuring the user privilege level directly on a user interface

To configure the user privilege level directly on a user interface that uses the scheme authentication mode:

Step	Command	Remarks
1. Configure the authentication type for SSH users as publickey .	For more information, see <i>Security Configuration Guide</i> .	Required only for SSH users who use public-key authentication.
2. Enter system view.	system-view	N/A
3. Enter user interface view.	user-interface { <i>first-num1</i> [<i>last-num1</i>] vtty <i>first-num2</i> [<i>last-num2</i>] }	N/A
4. Enable the scheme authentication mode.	authentication-mode scheme	By default, the authentication mode for VTY users is password , and no authentication is needed for AUX users.
5. Configure the user privilege level.	user privilege level <i>level</i>	By default, the user privilege level for users logged in through the AUX user interface is 3, and that for users logged in through the other user interfaces is 0.

To configure the user privilege level directly on a user interface that uses the **none** or **password** authentication mode:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter user interface view.	user-interface { <i>first-num1</i> [<i>last-num1</i>] { aux vtty } <i>first-num2</i> [<i>last-num2</i>] }	N/A
3. Configure the authentication mode for any user who uses the current user interface to log in to the device.	authentication-mode { none password }	Optional. By default, the authentication mode for VTY user interfaces is password , and no authentication is needed for AUX users.
4. Configure the privilege level of users logged in through the current user interface.	user privilege level <i>level</i>	Optional. By default, the user privilege level for users logged in through the AUX user interface is 3, and that for users logged in through the other user interfaces is 0.

For example:

```
# Display the commands a Telnet user can use by default after login.
```

```
<Sysname> ?
```

```
User view commands:
```

```
display Display current system information
```

```

ping      Ping function
quit      Exit from current command view
rsh       Establish one RSH connection
ssh2      Establish a secure shell client connection
super     Set the current user priority level
telnet    Establish one TELNET connection
tftp      Open TFTP connection
tracert   Trace route function

```

Configure the device to perform no authentication for Telnet users, and to authorize authenticated Telnet users to use level-0 and level-1 commands. (Use no authentication mode only in a secure network environment.)

```

<Sysname> system-view
[Sysname] user-interface vty 0 15
[Sysname-ui-vty0-15] authentication-mode none
[Sysname-ui-vty0-15] user privilege level 1

```

Display the commands a Telnet user can use after login. Because the user privilege level is 1, a Telnet user can use more commands now.

```
<Sysname> ?
```

User view commands:

```

debugging      Enable system debugging functions
dialer         Dialer disconnect
display        Display current system information
ping           Ping function
quit           Exit from current command view
refresh        Do soft reset
reset          Reset operation
rsh            Establish one RSH connection
screen-length  Specify the lines displayed on one screen
send           Send information to other user terminal interface
ssh2           Establish a secure shell client connection
super          Set the current user priority level
telnet         Establish one TELNET connection
terminal       Set the terminal line characteristics
tftp           Open TFTP connection
tracert        Trace route function
undo           Cancel current setting

```

Configure the device to perform password authentication for Telnet users, and to authorize authenticated Telnet users to use the commands of privilege levels 0, 1, and 2.

```

<Sysname> system-view
[Sysname] user-interface vty 0 15
[Sysname-ui-vty0-15] authentication-mode password
[Sysname-ui-vty0-15] set authentication password simple 123
[Sysname-ui-vty0-15] user privilege level 2

```

After the configuration is complete, when users Telnet to the device, they must enter the password **12345678**. After passing authentication, they can use commands of levels 0, 1, and 2.

Switching the user privilege level

Users can switch to a different user privilege level without logging out and terminating the current connection. After the privilege level switching, users can continue to manage the device without relogging in, but the commands they can execute have changed. For example, with the user privilege level 3, a user can configure system parameters. After switching to user privilege level 0, the user can execute only basic commands like **ping** and **tracert** and use a few **display** commands. The switching operation is effective for the current login. After the user relogs in, the user privilege restores to the original level.

To avoid problems, HP recommends that administrators log in with a lower privilege level to view switch operating parameters, and switch to a higher level temporarily only when they must maintain the device.

When an administrator must leave for a while or ask someone else to manage the device temporarily, they can switch to a lower privilege level before they leave to restrict the operation by others.

Configuring the authentication parameters for user privilege level switching

A user can switch to a privilege level equal to or lower than the current one unconditionally and is not required to enter a password (if any).

For security, a user is required to enter a password (if any) to switch to a higher privilege level. The authentication falls into one of the following categories:

Keywords	Authentication mode	Description
local	Local password authentication only (local-only)	The device authenticates a user by using the privilege level switching password entered by the user. To use this mode, you must set the password for privilege level switching by using the super password command.
scheme	Remote AAA authentication through HWTACACS or RADIUS	The device sends the username and password for privilege level switching to the HWTACACS or RADIUS server for remote authentication. To use this mode, you must perform the following configuration tasks: <ul style="list-style-type: none">• Configure the required HWTACACS or RADIUS schemes and configure the ISP domain to use the schemes for users. For more information, see <i>Security Configuration Guide</i>.• Add user accounts and specify the user passwords on the HWTACACS or RADIUS server.
local scheme	Local password authentication first and then remote AAA authentication	The device authenticates a user by using the local password first, and if no password for privilege level switching is set, for the user logged in to the AUX user interface, the privilege level is switched directly; for VTY users, AAA authentication is performed.
scheme local	Remote AAA authentication first and then local password authentication	AAA authentication is performed first, and if the remote HWTACACS or RADIUS server does not respond or AAA configuration on the device is invalid, the local password authentication is performed.

To configure the authentication parameters for a user privilege level:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the authentication mode for user privilege level switching.	super authentication-mode { local scheme } *	Optional. By default, local-only authentication is used.
3. Configure the password for a user privilege level.	super password [level <i>user-level</i>] { cipher simple } <i>password</i>	Required for local authentication. By default, a privilege level has no password. If no user privilege level is specified when you configure the command, the user privilege level defaults to 3.

If local-only authentication is used, an AUX user interface user (a user logged in through the console port) can switch to a higher privilege level even if the privilege level has not been assigned a password.

Switching to a higher user privilege level

Before you switch to a higher user privilege level, obtain the required authentication data as described in Table 8.

The privilege level switching fails after three consecutive unsuccessful password attempts.

To switch the user privilege level, perform the following task in user view:

Task	Command	Remarks
Switch the user privilege level.	super [<i>level</i>]	When logging in to the device, a user has a user privilege level, which depends on user interface or authentication user level.

Table 8 Information required for user privilege level switching

User interface authentication mode	User privilege level switching authentication mode	Information required for the first authentication mode	Information required for the second authentication mode
none/password	local	Password configured on the device with the super password command for the privilege level	N/A
	local scheme	Password configured on the device with the super password command for the privilege level	Username and password configured on the AAA server for the privilege level
	scheme	Username and password for the privilege level	N/A
	scheme local	Username and password for the privilege level	Local user privilege level switching password

User interface authentication mode	User privilege level switching authentication mode	Information required for the first authentication mode	Information required for the second authentication mode
scheme	local	Password configured on the device with the super password command for the privilege level	N/A
	local scheme	Password configured on the device with the super password command for the privilege level	Password for privilege level switching (configured on the AAA server). The system uses the username used for logging in as the privilege level switching username.
	scheme	Password for privilege level switching (configured on the AAA server). The system uses the username used for logging in as the privilege level switching username.	N/A
	scheme local	Password for privilege level switching (configured on the AAA server). The system uses the username used for logging in as the privilege level switching username.	Password configured on the device with the super password command for the privilege level

Changing the level of a command

Every command in a view has a default command level. The default command level scheme is sufficient for the security and ease of maintenance requirements of most networks. If you want to change the level of a command, make sure the change does not result in any security risk or maintenance problem.

To change the level of a command:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Change the level of a command in a specific view.	command-privilege level level view view command	See Table 7 for the default settings.

Saving the running configuration

You can use the **save** command in any view to save all submitted and executed commands into the configuration file. Commands saved in the configuration file can survive a reboot. The **save** command does not take effect on one-time commands, including **display** and **reset** commands. One-time commands are never saved.

Displaying and maintaining CLI

Task	Command	Remarks
Display the command keyword alias configuration.	display command-alias [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display data in the clipboard.	display clipboard [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Login overview

This chapter describes the available CLI login methods and their configuration procedures.

Login methods at a glance

You can access the device only through the console port at the first login, locally or remotely by using a pair of modems. After you log in to the device, you can configure other login methods, including Telnet and SSH, for remote access.

Table 9 Login methods

Login method	Default setting and configuration requirements
Logging in to the CLI:	
<ul style="list-style-type: none">Logging in through the console port for the first time	By default, login through the console port is enabled, no username or password is required, and the user privilege level is 3.
<ul style="list-style-type: none">Logging in through Telnet	By default, Telnet service is enabled. To use Telnet service, complete the following configuration tasks: <ul style="list-style-type: none">Enable the Telnet server.Assign an IP address to a Layer 3 interface and make sure the interface and the Telnet client can reach each other.Configure the authentication mode for VTY login users (password by default).Configure the user privilege level of VTY login users (0 by default).
<ul style="list-style-type: none">Logging in through SSH	By default, SSH service is disabled. To use SSH service, complete the following configuration tasks: <ul style="list-style-type: none">Enable the SSH function and configure SSH attributes.Assign an IP address to a Layer 3 interface and make sure the interface and the SSH client can reach each other.Enable scheme authentication for VTY login users.Configure the user privilege level of VTY login users (0 by default).
<ul style="list-style-type: none">Modem dial-in through the console port	By default, modem dial-in is enabled, no username or password is required, and the user privilege level is 3.
Logging in to the Web interface	
	By default, Web login is disabled. To use Web service, complete the following configuration tasks: <ul style="list-style-type: none">Assign an IP address to a Layer 3 interface.Configure a local user account for Web login, and assign a user privilege level and the Web service to the account.
Logging in through NMS	
	By default, SNMP login is disabled. To use SNMP service, complete the following configuration tasks: <ul style="list-style-type: none">Assign an IP address to a Layer 3 interface, and make sure the interface and the NMS can reach each other.Configure SNMP basic parameters.

User interfaces

The device uses user interfaces (also called "lines") to control CLI logins and monitor CLI sessions. You can configure access control settings, including authentication, user privilege, and login redirect on user interfaces. After users are logged in, their actions must be compliant with the settings on the user interfaces assigned to them.

Users are assigned different user interfaces, depending on their login methods, as shown in Table 10.

Table 10 CLI login method and user interface matrix

User interface	Login method
AUX user interface	Console port (EIA/TIA-232 DCE), locally or remotely by using modems
Virtual type terminal (VTY) user interface	Telnet or SSH

User interface assignment

The device automatically assigns user interfaces to CLI login users, depending on their login methods. Each user interface can be assigned to only one user at a time. If no user interface is available, a CLI login attempt will be rejected.

The device provides one AUX user interfaces and 16 VTY user interfaces. For a CLI login, the device always picks the lowest numbered user interface from the idle user interfaces available for the type of login.

For example, four VTY user interfaces (0 to 3) are configured, of which VTY 0 and VTY 3 are idle. When a user Telnets to the device, the device assigns VTY 0 to the user and uses the settings on VTY 0 to authenticate and manage the user.

User interface numbering

User interfaces are numbered by using absolute numbering or relative numbering.

Absolute numbering

An absolute number uniquely identifies a user interface among all user interfaces. The user interfaces are numbered starting from 0 and incrementing by 1 and in the sequence of AUX and VTY user interfaces. You can use the **display user-interface** command without any parameters to view supported user interfaces and their absolute numbers.

Relative numbering

A relative number uniquely identifies a user interface among all user interfaces that are the same type. The number format is *user interface type + number*. All the types of user interfaces are numbered starting from 0 and incrementing by 1. For example, the first AUX user interface is AUX 0.

A relative number uniquely identifies a user interface among all user interfaces that are the same type. The number format is *user interface type + number*. The user interfaces are numbered starting from 0 and incrementing by 1. For example, the first AUX user interface is AUX 0, and the second AUX user interface is AUX 1.

Logging in to the CLI

By default, the first time you access the CLI you must log in through the console port, locally or remotely by using a pair of modems. At the CLI, you can configure Telnet or SSH for remote access.

Logging in through the console port for the first time

To log in through the console port, make sure the console terminal has a terminal emulation program (for example, HyperTerminal in Windows XP). In addition, the port settings of the terminal emulation program must be the same as the default settings of the console port in Table 11.

Table 11 Default console port properties

Parameter	Default
Bits per second	9600 bps
Flow control	None
Parity	None
Stop bits	1
Data bits	8

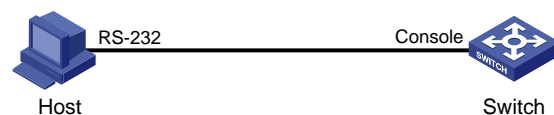
To log in through the console port from a console terminal (for example, a PC):

1. Plug the DB-9 female connector of the console cable to the serial port of the PC.
2. Plug the RJ-45 connector of the console cable to the console port of the device.

NOTE:

- Identify the mark on the console port and make sure you are connecting to the correct port.
- The serial ports on PCs do not support hot swapping. If the switch has been powered on, always connect the console cable to the PC before connecting to the switch, and when you disconnect the cable, first disconnect it from the switch.

Figure 4 Connecting a terminal to the console port



3. If the PC is off, turn on the PC.

Launch the terminal emulation program and configure the communication properties on the PC. Figure 5 through Figure 7 show the configuration procedure on Windows XP HyperTerminal. Make sure the port settings are the same as listed in Table 11.

NOTE:

On Windows Server 2003, add the HyperTerminal program first, and then log in to and manage the device as described in this document. On Windows Server 2008, Windows 7, Windows Vista, or some other operating system, obtain a third-party terminal control program first, and then follow the user guide or online help to log in to the device.

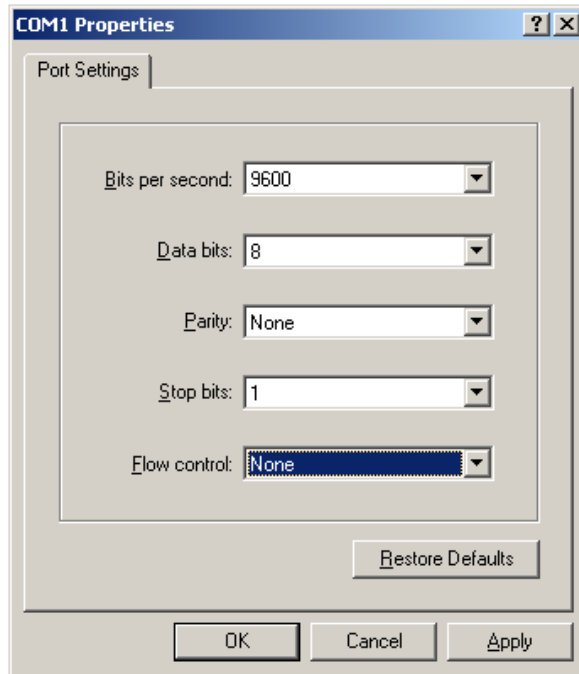
Figure 5 Connection description



Figure 6 Specifying the serial port used to establish the connection

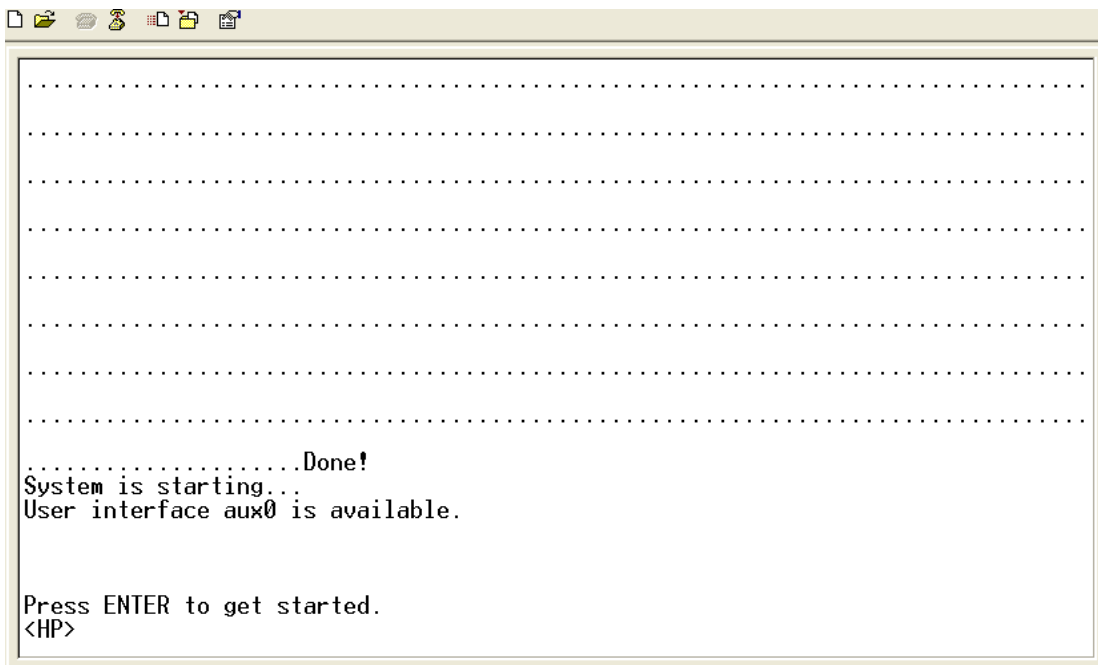


Figure 7 Setting the properties of the serial port



4. Power on the device and press **Enter** at the prompt.

Figure 8 CLI



5. At the default user view prompt <HP>, enter commands to configure the device or view the running status of the device. To get help, enter ?.

Configuring console login control settings

The following authentication modes are available for controlling console logins:

- **None**—Requires no authentication. This mode is insecure.
- **Password**—Requires password authentication. If your password was lost, see *HP Series Ethernet Switches Login Password Recovery Manual* for password recovery.
- **Scheme**—Uses the AAA module to provide local or remote console login authentication. You must provide a username and password for accessing the CLI. If the password configured in the local user database was lost, see *HP Series Ethernet Switches Login Password Recovery Manual* for password recovery. If the username or password configured on a remote server was lost, contact the server administrator for help.

By default, console login does not require authentication. Any user can log in through the console port without authentication and have user privilege level 3. To improve device security, configure the password or scheme authentication mode immediately after you log in to the device for the first time.

Table 12 Configuration required for different console login authentication modes

Authentication mode	Configuration tasks	Reference
None	Set the authentication mode to none for the AUX user interface.	"Configuring none authentication for console login"
Password	Enable password authentication on the AUX user interface. Set a password.	"Configuring password authentication for console login"
Scheme	Enable scheme authentication on the AUX user interface. Configure local or remote authentication settings. To configure local authentication: <ol style="list-style-type: none"> 3. Configure a local user and specify the password. 4. Configure the device to use local authentication. To configure remote authentication: <ol style="list-style-type: none"> 5. Configure the RADIUS or HWTACACS scheme on the device. Configure the username and password on the AAA server. 6. Configure the device to use the scheme for user authentication. 	"Configuring scheme authentication for console login"

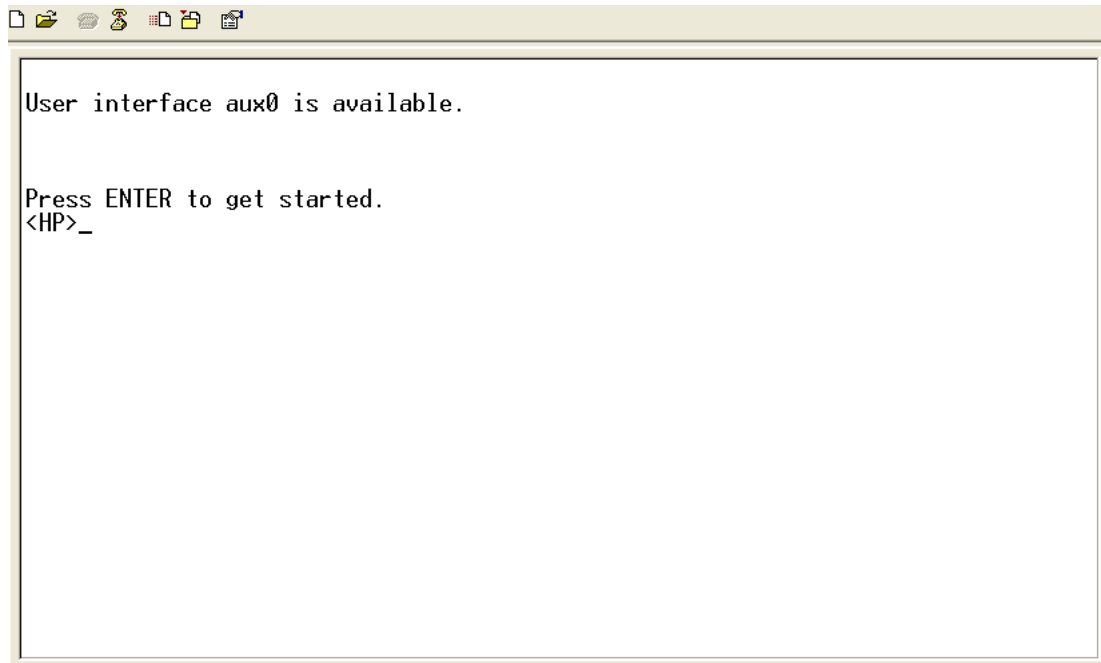
Configuring none authentication for console login

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter AUX user interface view.	user-interface aux <i>first-number</i> [<i>last-number</i>]	N/A
3. Enable the none authentication mode.	authentication-mode none	By default, you can log in to the device through the console port without authentication and have user privilege level 3.

Step	Command	Remarks
4.	Configure common settings for console login.	See "Configuring common console login settings (optional)." Optional.

The next time you attempt to log in through the console port, you do not need to provide any username or password, as shown in [Figure 9](#).

Figure 9 Accessing the CLI through the console port without authentication

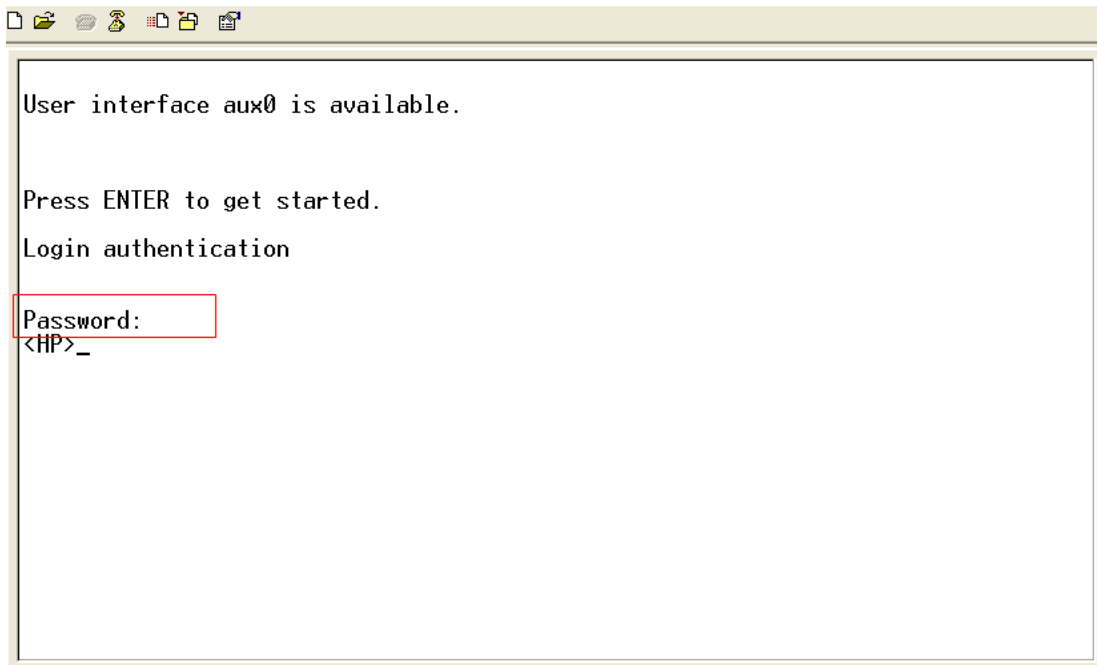


Configuring password authentication for console login

Step	Command	Remarks
1.	Enter system view.	system-view N/A
2.	Enter AUX user interface view.	user-interface aux <i>first-number</i> [<i>last-number</i>] N/A
3.	Enable password authentication.	authentication-mode password By default, you can log in to the device through the console port without authentication and have user privilege level 3 after login.
4.	Set a password.	set authentication password { cipher simple } <i>password</i> By default, no password is set.
5.	Configure common settings for console login.	See "Configuring common console login settings (optional)." Optional.

The next time you attempt to log in through the console port, you must provide the configured login password, as shown in [Figure 10](#).

Figure 10 Password authentication interface for console login



Configuring scheme authentication for console login

Follow these guidelines when you configure scheme authentication for console login:

- To make the command authorization or command accounting function take effect, apply an HWTACACS scheme to the intended ISP domain. This scheme must specify the IP address of the authorization server and other authorization parameters.
- If the local authentication scheme is used, use the **authorization-attribute level** *level* command in local user view to set the user privilege level on the device.
- If a RADIUS or HWTACACS authentication scheme is used, set the user privilege level on the RADIUS or HWTACACS server.

To configure scheme authentication for console login:

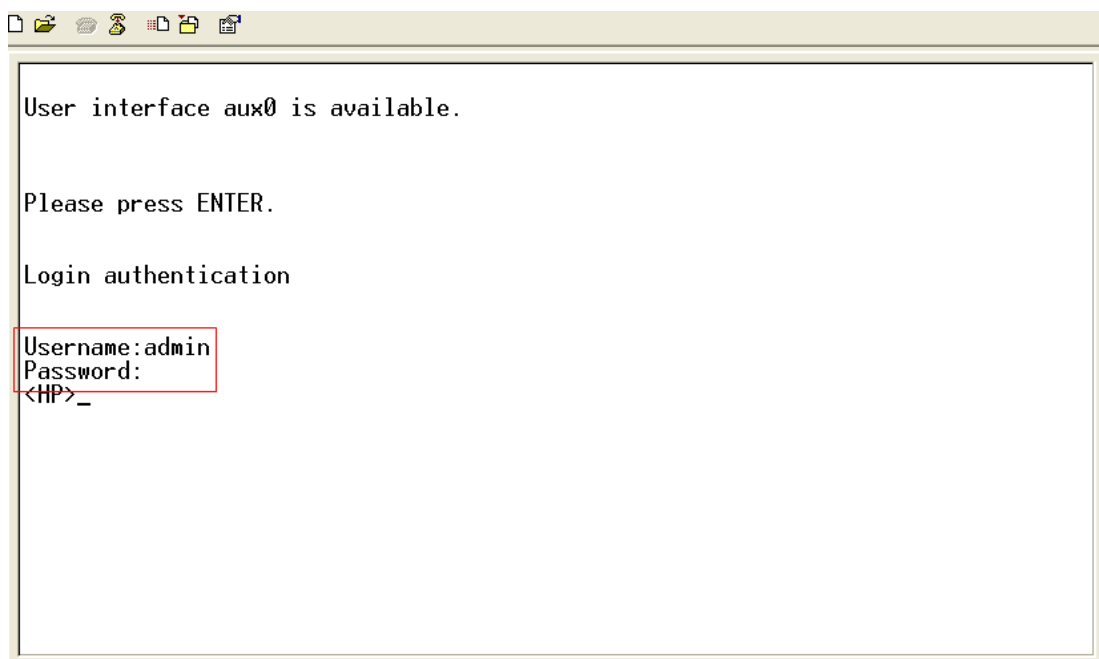
Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter AUX user interface view.	user-interface aux <i>first-number</i> [<i>last-number</i>]	N/A
3. Enable scheme authentication.	authentication-mode scheme	Whether local, RADIUS, or HWTACACS authentication is adopted depends on the configured AAA scheme. By default, console log users are not authenticated.

Step	Command	Remarks
4. Enable command authorization.	command authorization	Optional. By default, command authorization is disabled. The commands available for a user only depend on the user privilege level. If command authorization is enabled, a command is available only if the user has the commensurate user privilege level and is authorized to use the command by the AAA scheme.
5. Enable command accounting.	command accounting	Optional. By default, command accounting is disabled. The accounting server does not record the commands executed by users. Command accounting allows the HWTACACS server to record all commands executed by users, regardless of command execution results. This function helps control and monitor user behaviors on the device. If command accounting is enabled and command authorization is not enabled, every executed command is recorded on the HWTACACS server. If both command accounting and command authorization are enabled, only the authorized and executed commands are recorded on the HWTACACS server.
6. Exit to system view.	quit	N/A
7. Apply an AAA authentication scheme to the intended domain.	<p>8. Enter ISP domain view: domain <i>domain-name</i></p> <p>9. Apply an AAA scheme to the domain: authentication default { hwtacacs-scheme <i>hwtacacs-scheme-name</i> [local] local none radius-scheme <i>radius-scheme-name</i> [local] }</p> <p>10. Exit to system view: quit</p>	Optional. By default, local authentication is used. For local authentication, configure local user accounts. For RADIUS or HWTACACS authentication, configure the RADIUS or HWTACACS scheme on the device and configure authentication settings (including the username and password) on the server. For more information about AAA configuration, see <i>Security Configuration Guide</i> .
11. Create a local user and enter local user view.	local-user <i>user-name</i>	By default, no local user exists.

Step	Command	Remarks
12. Set an authentication password for the local user.	password { cipher simple } <i>password</i>	By default, no password is set.
13. Specifies a command level of the local user.	authorization-attribute level <i>level</i>	Optional. By default, the command level is 0.
14. Specify terminal service for the local user.	service-type terminal	By default, no service type is specified.
15. Configure common settings for console login.	See " Configuring common console login settings (optional) ."	Optional.

The next time you attempt to log in through the console port, you must provide the configured login username and password, as shown in [Figure 11](#).

Figure 11 Scheme authentication interface for console login



Configuring common console login settings (optional)

Some common settings configured for an AUX user interface take effect immediately and can interrupt the console login session. To save you the trouble of repeated re-logins, use a login method different from console login to log in to the device before you change console login settings.

After the configuration is complete, change the terminal settings on the configuration terminal and make sure they are the same as the settings on the device.

To configure common settings for an AUX user interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable copyright information display.	copyright-info enable	By default, copyright information display is enabled.

Step	Command	Remarks
3. Enter AUX user interface view.	user-interface aux <i>first-number</i> [<i>last-number</i>]	N/A
4. Configure the baud rate.	speed <i>speed-value</i>	By default, the transmission rate is 9600 bps.
5. Configure the parity check mode.	parity { even none odd }	The default setting is none , namely, no parity check.
6. Configure the number of stop bits.	stopbits { 1 1.5 2 }	The default is 1. Stop bits indicate the end of a character. The more the stop bits, the slower the transmission.
7. Configure the number of data bits in a character.	databits { 7 8 }	By default, the number of data bits in each character is 8. The setting depends on the character coding type. For example, you can set it to 7 if standard ASCII characters are to be sent, and set it to 8 if extended ASCII characters are to be sent.
8. Define a shortcut key for enabling a terminal session.	activation-key <i>character</i>	By default, press Enter to enable a terminal session.
9. Define a shortcut key for terminating tasks.	escape-key { default <i>character</i> }	By default, press Ctrl+C to terminate a task.
10. Configure the flow control mode.	flow-control { hardware none software }	By default, the flow control mode is none . The device supports only the none mode.
11. Specify the terminal display.	terminal type { ansi vt100 }	By default, the terminal display type is ANSI. The device supports two terminal display types: ANSI and VT100. HP recommends setting the display type to VT100 for both the device and the client. If the device and the client use different display types or both use the ANSI display type, when the total number of characters of a command line exceeds 80, the screen display on the terminal might be abnormal. For example, the cursor might be displayed at a wrong place.
12. Configure the user privilege level for login users.	user privilege level <i>level</i>	By default, the default command level is 3 for AUX user interfaces.
13. Set the maximum number of lines to be displayed on a screen.	screen-length <i>screen-length</i>	By default, a screen displays 24 lines at most. A value of 0 disables pausing between screens of output.

Step	Command	Remarks
14. Set the size of command history buffer.	history-command max-size <i>value</i>	By default, the buffer saves 10 history commands at most.
15. Set the idle-timeout timer.	idle-timeout <i>minutes</i> [<i>seconds</i>]	The default idle-timeout is 10 minutes. The system automatically terminates the user's connection if there is no information interaction between the device and the user within the idle-timeout time. Setting idle-timeout to 0 disables the timer.

Logging in through Telnet

You can Telnet to the device through a VTY user interface for remote management, or use the device as a Telnet client to Telnet to other devices, as shown in [Figure 12](#).

Figure 12 Telnet login



Table 13 shows the Telnet server and client configuration required for a successful Telnet login.

Table 13 Telnet server and Telnet client configuration requirements

Object	Requirements
Telnet server	Enable Telnet server Assign an IP address to a Layer 3 interface, and make sure the Telnet server and client can reach each other. Configure the authentication mode and other settings.
Telnet client	Run the Telnet client program. Obtain the IP address of the Layer 3 interface on the server.

To control Telnet access to the device working as a Telnet server, configure authentication and user privilege for Telnet users.

By default, password authentication applies to Telnet login, but no login password is configured. To allow Telnet access to the device after you enable the Telnet server, you must configure a password.

The following are authentication modes available for controlling Telnet logins:

- **None**—Requires no authentication and is insecure.
- **Password**—Requires a password for accessing the CLI. If your password was lost, log in to the device through the console port to modify the password.
- **Scheme**—Uses the AAA module to provide local or remote authentication. You must provide a username and password for accessing the CLI. If the password configured in the local user database was lost, see *HP Series Ethernet Switches Login Password Recovery Manual* for password

recovery. If the username or password configured on a remote server was lost, contact the server administrator for help.

Table 14 Configuration required for different Telnet login authentication modes

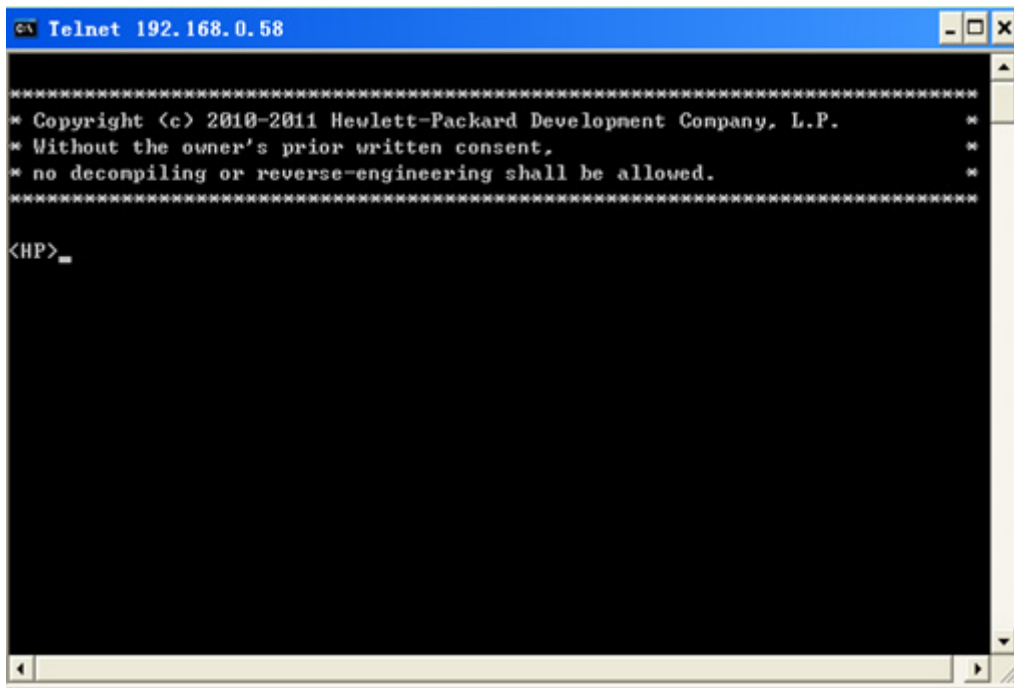
Authentication mode	Configuration tasks	Reference
None	Set the authentication mode to none for the VTY user interface.	"Configuring none authentication for Telnet login"
Password	Enable password authentication on the VTY user interface. Set a password.	"Configuring password authentication for Telnet login"
AAA	Enable scheme authentication on the VTY user interface. Configure local or remote authentication settings. To configure local authentication: <ol style="list-style-type: none"> 1. Configure a local user and specify the password. 2. Configure the device to use local authentication. To configure remote authentication: <ol style="list-style-type: none"> 3. Configure the RADIUS or HWTACACS scheme on the device. 4. Configure the username and password on the AAA server. 5. Configure the device to use the scheme for user authentication. 	"Configuring scheme authentication for Telnet login"

Configuring none authentication for Telnet login

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable Telnet server.	telnet server enable	By default, the Telnet server is disabled.
3. Enter one or multiple VTY user interface views.	user-interface vty <i>first-number</i> [<i>last-number</i>]	N/A
4. Enable the none authentication mode.	authentication-mode none	By default, authentication mode for VTY user interfaces is password .
5. Configure the command level for login users on the current user interfaces.	user privilege level <i>level</i>	By default, the default command level is 0 for VTY user interfaces.
6. Configure common settings for the VTY user interfaces.	See "Configuring common settings for VTY user interfaces (optional)."	Optional.

The next time you attempt to Telnet to the device, you do not need to provide any username or password, as shown in [Figure 13](#). If the maximum number of login users has been reached, your login attempt fails and the message "All user interfaces are used, please try later!" appears.

Figure 13 Telnetting to the device without authentication

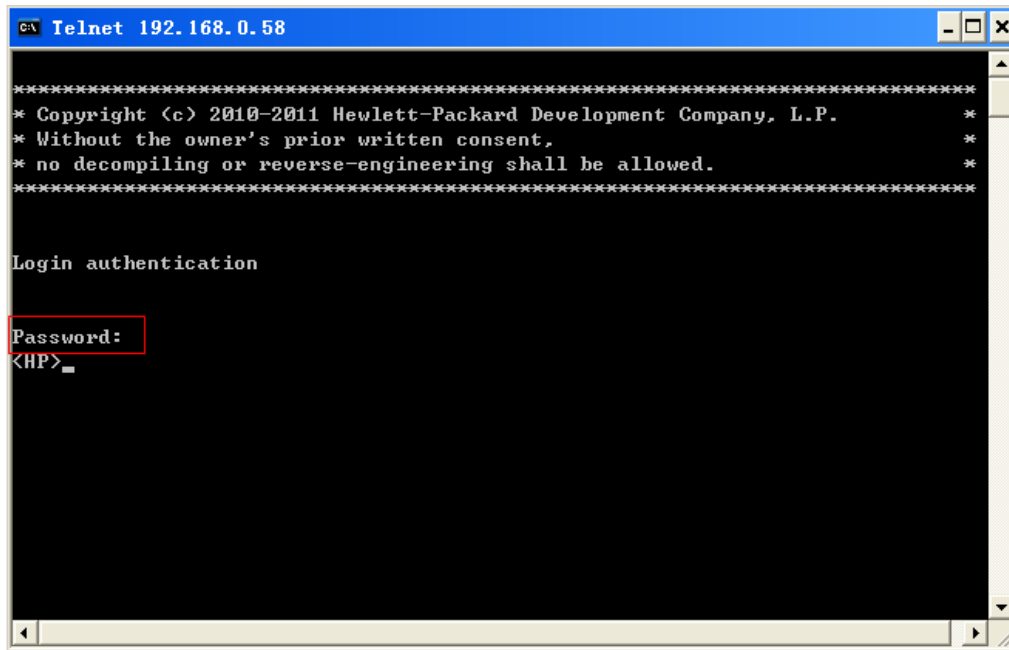


Configuring password authentication for Telnet login

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable Telnet.	telnet server enable	By default, the Telnet service is disabled.
3. Enter one or multiple VTY user interface views.	user-interface vty <i>first-number</i> [<i>last-number</i>]	N/A
4. Enable password authentication.	authentication-mode password	By default, password authentication is enabled for VTY user interfaces.
5. Set a password.	set authentication password { cipher simple } <i>password</i>	By default, no password is set.
6. Configure the user privilege level for login users.	user privilege level <i>level</i>	The default level is 0.
7. Configure common settings for VTY user interfaces.	See " Configuring common settings for VTY user interfaces (optional) ."	Optional.

The next time you attempt to Telnet to the device, you must provide the configured login password, as shown in [Figure 14](#). If the maximum number of login users has been reached, your login attempt fails and the message "All user interfaces are used, please try later!" appears.

Figure 14 Password authentication interface for Telnet login



Configuring scheme authentication for Telnet login

Follow these guidelines when you configure scheme authentication for Telnet login:

- To make the command authorization or command accounting function take effect, apply an HWTACACS scheme to the intended ISP domain. This scheme must specify the IP address of the authorization server and other authorization parameters.
- If the local authentication scheme is used, use the **authorization-attribute level** *level* command in local user view to set the user privilege level on the device.
- If a RADIUS or HWTACACS authentication scheme is used, set the user privilege level on the RADIUS or HWTACACS server.

To configure scheme authentication for Telnet login:

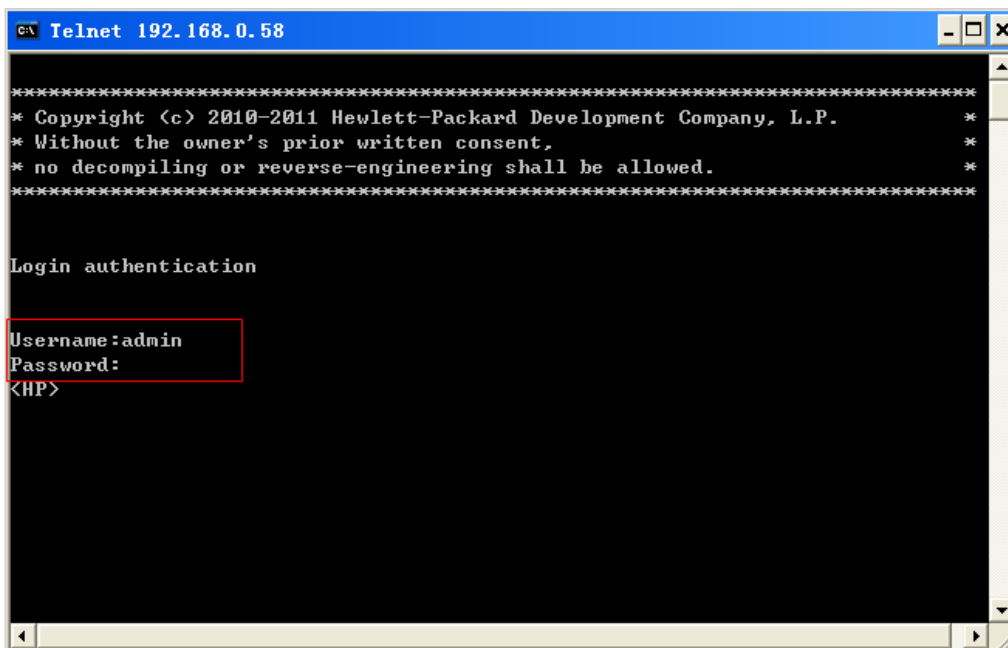
Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable Telnet.	telnet server enable	By default, the Telnet service is disabled.
3. Enter one or multiple VTY user interface views.	user-interface vty <i>first-number</i> [<i>last-number</i>]	N/A
4. Enable scheme authentication.	authentication-mode scheme	Whether local, RADIUS, or HWTACACS authentication is adopted depends on the configured AAA scheme. By default, local authentication is adopted.

Step	Command	Remarks
5. Enable command authorization.	command authorization	<p>Optional.</p> <p>By default, command authorization is disabled. The commands available for a user only depend on the user privilege level.</p> <p>If command authorization is enabled, a command is available only if the user has the commensurate user privilege level and is authorized to use the command by the AAA scheme.</p>
6. Enable command accounting.	command accounting	<p>Optional.</p> <p>By default, command accounting is disabled. The accounting server does not record the commands executed by users.</p> <p>Command accounting allows the HWTACACS server to record all executed commands that are supported by the device, regardless of the command execution result. This function helps control and monitor user behaviors on the device. If command accounting is enabled and command authorization is not enabled, every executed command is recorded on the HWTACACS server. If both command accounting and command authorization are enabled, only the authorized and executed commands are recorded on the HWTACACS server.</p>
7. Exit to system view.	quit	N/A
8. Apply an AAA authentication scheme to the intended domain.	<ol style="list-style-type: none"> 1. Enter ISP domain view: domain <i>domain-name</i> 2. Apply an AAA scheme to the domain: authentication default { hwtacacs-scheme <i>hwtacacs-scheme-name</i> [local] local none radius-scheme <i>radius-scheme-name</i> [local] } 3. Exit to system view: quit 	<p>Optional.</p> <p>By default, local authentication is used.</p> <p>For local authentication, configure local user accounts.</p> <p>For RADIUS or HWTACACS authentication, configure the RADIUS or HWTACACS scheme on the device and configure authentication settings (including the username and password) on the server.</p> <p>For more information about AAA configuration, see <i>Security Configuration Guide</i>.</p>

Step	Command	Remarks
9. Create a local user and enter local user view.	local-user <i>user-name</i>	By default, no local user exists.
10. Set a password.	password { cipher simple } <i>password</i>	By default, no password is set.
11. Specify the command level of the local user.	authorization-attribute level <i>level</i>	Optional. By default, the command level is 0.
12. Specify Telnet service for the local user.	service-type telnet	By default, no service type is specified.
13. Exit to system view.	quit	N/A
14. Configure common settings for VTY user interfaces.	See " Configuring common settings for VTY user interfaces (optional) ."	Optional.

The next time you attempt to Telnet to the CLI, you must provide the configured login username and password, as shown in [Figure 15](#). If you are required to pass a second authentication, you must also provide the correct password to access the CLI. If the maximum number of login users has been reached, your login attempt fails and the message "All user interfaces are used, please try later!" appears.

Figure 15 Scheme authentication interface for Telnet login



Configuring common settings for VTY user interfaces (optional)

You might be unable to access the CLI through a VTY user interface after configuring the **auto-execute command** command on it. Before you configure the command and save the configuration, make sure you can access the CLI through a different user interface.

To configure common settings for VTY user interfaces:

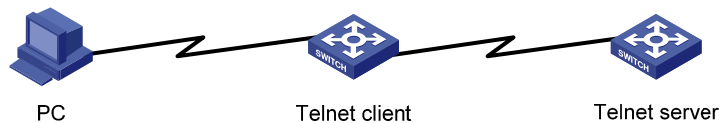
Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable copyright information display.	copyright-info enable	By default, copyright information display is enabled.
3. Enter one or multiple VTY user interface views.	user-interface vty <i>first-number</i> [<i>last-number</i>]	N/A
4. Enable the terminal service.	shell	Optional. By default, terminal service is enabled.
5. Enable the user interfaces to support Telnet, SSH, or both of them.	protocol inbound { all ssh telnet }	Optional. By default, both Telnet and SSH are supported. The configuration takes effect the next time you log in.
6. Define a shortcut key for terminating tasks.	escape-key { default <i>character</i> }	Optional. By default, press Ctrl+C to terminate a task.
7. Configure the type of terminal display.	terminal type { ansi vt100 }	Optional. By default, the terminal display type is ANSI.
8. Set the maximum number of lines to be displayed on a screen.	screen-length <i>screen-length</i>	Optional. By default, a screen displays 24 lines. A value of 0 disables the function.
9. Set the size of command history buffer.	history-command max-size <i>value</i>	Optional. By default, the buffer saves 10 history commands.
10. Set the idle-timeout timer.	idle-timeout <i>minutes</i> [<i>seconds</i>]	Optional. The default idle-timeout is 10 minutes for all user interfaces. The system automatically terminates the user's connection if there is no information interaction between the device and the user within the timeout time. Setting idle-timeout to 0 disables the timer.

Step	Command	Remarks
11. Specify a command to be automatically executed when a user logs in to the user interfaces.	auto-execute command <i>command</i>	Optional. By default, no automatically executed command is specified. The command auto-execute function is typically used for redirecting a Telnet user to a specific host. After executing the specified command and performing the incurred task, the system automatically disconnect the Telnet session.

Using the device to log in to a Telnet server

You can use the device as a Telnet client to log in to a Telnet server. If the server is located in a different subnet than the device, make sure the two devices have routes to reach each other.

Figure 16 Telnetting from the device to a Telnet server



To use the device to log in to a Telnet server:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Specify a source IPv4 address or source interface for outgoing Telnet packets.	telnet client source { interface <i>interface-type interface-number</i> ip <i>ip-address</i> }	Optional. By default, no source IPv4 address or source interface is specified. The IP address of the outbound interface is used as the source IPv4 address.
3. Exit to user view.	quit	N/A
4. Use the device to log in to a Telnet server.	<ul style="list-style-type: none"> Log in to an IPv4 Telnet server: telnet <i>remote-host</i> [<i>service-port</i>] [[vpn-instance <i>vpn-instance-name</i>] [source { interface <i>interface-type</i> <i>interface-number</i> ip <i>ip-address</i> }]] Log in to an IPv6 Telnet server: telnet ipv6 <i>remote-host</i> [-i <i>interface-type</i> <i>interface-number</i>] [<i>port-number</i>] [vpn-instance <i>vpn-instance-name</i>] 	Use either command. The vpn-instance <i>vpn-instance-name</i> option is only available on the HP 5500-EI switches.

Setting the DSCP value for IP to use for outgoing Telnet packets

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the DSCP value for IP to use for outgoing Telnet packets.	<ul style="list-style-type: none"> On a Telnet client running IPv4: telnet client dscp dscp-value On a Telnet client running IPv6: telnet client ipv6 dscp dscp-value On a Telnet server running IPv4: telnet server dscp dscp-value On a Telnet server running IPv6: telnet server ipv6 dscp dscp-value 	<p>The default is as follows:</p> <ul style="list-style-type: none"> 16 for a Telnet client running IPv4. 0 for a Telnet client running IPv6. 48 for a Telnet server running IPv4. 0 for a Telnet server running IPv6.

Logging in through SSH

SSH offers a secure approach to remote login. By providing encryption and strong authentication, it protects devices against attacks such as IP spoofing and plaintext password interception. You can log in to the device working as an SSH server for remote management, as shown in Figure 17. You can also use the device as an SSH client to log in to an SSH server.

Figure 17 SSH login diagram



Table 15 shows the SSH server and client configuration required for a successful SSH login.

Table 15 SSH server and client requirements

Device role	Requirements
SSH server	<p>Assign an IP address to a Layer 3 interface, and make sure the interface and the client can reach each other.</p> <p>Configure the authentication mode and other settings.</p>
SSH client	<p>If the host operates as an SSH client, run the SSH client program on the host.</p> <p>Obtain the IP address of the Layer 3 interface on the server.</p>

To control SSH access to the device working as an SSH server, configure authentication and user privilege level for SSH users.

By default, password authentication is adopted for SSH login, but no login password is configured. To allow SSH access to the device after you enable the SSH server, you must configure a password.

Configuring the SSH server on the device

Follow these guidelines when you configure the SSH server:

- To make the command authorization or command accounting function take effect, apply an HWTACACS scheme to the intended ISP domain. This scheme must specify the IP address of the authorization server and other authorization parameters.
- If the local authentication scheme is used, use the **authorization-attribute level** *level* command in local user view to set the user privilege level on the device.
- If a RADIUS or HWTACACS authentication scheme is used, set the user privilege level on the RADIUS or HWTACACS server.

The SSH client authentication method is password in this configuration procedure. For more information about SSH and publickey authentication, see *Security Configuration Guide*.

To configure the SSH server on the device:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create local key pairs.	public-key local create { dsa rsa }	By default, no local key pairs are created.
3. Enable SSH server.	ssh server enable	By default, SSH server is disabled.
4. Enter one or more VTY user interface views.	user-interface vty <i>first-number</i> [<i>last-number</i>]	N/A
5. Enable scheme authentication.	authentication-mode scheme	By default, password authentication is enabled on VTY user interfaces.
6. Enable the user interfaces to support Telnet, SSH, or both of them.	protocol inbound { all ssh Telnet }	Optional. By default, both Telnet and SSH are supported.
7. Enable command authorization.	command authorization	Optional. By default, command authorization is disabled. The commands available for a user only depend on the user privilege level. If command authorization is enabled, a command is available only if the user has the commensurate user privilege level and is authorized to use the command by the AAA scheme.

Step	Command	Remarks
8. Enable command accounting.	command accounting	Optional. By default, command accounting is disabled. The accounting server does not record the commands executed by users. Command accounting allows the HWTACACS server to record all executed commands that are supported by the device, regardless of the command execution result. This function helps control and monitor user behaviors on the device. If command accounting is enabled and command authorization is not enabled, every executed command is recorded on the HWTACACS server. If both command accounting and command authorization are enabled, only the authorized and executed commands are recorded on the HWTACACS server.
9. Exit to system view.	quit	N/A
10. Apply an AAA authentication scheme to the intended domain.	<ol style="list-style-type: none"> Enter the ISP domain view: domain <i>domain-name</i> Apply the specified AAA scheme to the domain: authentication default { hwtacacs-scheme <i>hwtacacs-scheme-name</i> [local] local none radius-scheme <i>radius-scheme-name</i> [local] } Exit to system view: quit 	Optional. For local authentication, configure local user accounts. For RADIUS or HWTACACS authentication, configure the RADIUS or HWTACACS scheme on the device and configure authentication settings (including the username and password) on the server. For more information about AAA configuration, see <i>Security Configuration Guide</i> .
11. Create a local user and enter local user view.	local-user <i>user-name</i>	By default, no local user exists.
12. Set a password for the local user.	password { cipher simple } <i>password</i>	By default, no password is set.
13. Specify the command level of the user.	authorization-attribute level <i>level</i>	Optional. By default, the command level is 0.
14. Specify SSH service for the user.	service-type <i>ssh</i>	By default, no service type is specified.
15. Exit to system view.	quit	N/A

Step	Command	Remarks
16. Create an SSH user, and specify the authentication mode for the SSH user.	ssh user <i>username</i> service-type stelnet authentication-type { password { any password-publickey publickey } assign publickey <i>keyname</i> }	N/A By default, no SSH user is created.
17. Configure common settings for VTY user interfaces.	See " Configuring common settings for VTY user interfaces (optional) ."	Optional.

Using the device as an SSH client to log in to the SSH server

You can use the device as an SSH client to log in to an SSH server. If the server is located in a different subnet than the device, make sure the two devices have routes to reach each other.

Figure 18 Logging in to an SSH server from the device



To use the device as an SSH client to log in to an SSH server, perform the following tasks in user view:

Task	Command	Remarks
Log in to an IPv4 SSH server.	ssh2 server	The <i>server</i> argument represents the IPv4 address or host name of the server.
Log in to an IPv6 SSH server.	ssh2 ipv6 server	The <i>server</i> argument represents the IPv6 address or host name of the server.

To work with the SSH server, you might need to configure the SSH client. For information about configuring the SSH client, see *Security Configuration Guide*.

Modem dial-in through the console port

You can use a pair of modems to remotely connect to a device through its console port over the PSTN when the IP network connection is broken. To do so, make sure the dial-in connection, the device, and the modems are correctly set up.

By default, you can log in to the device through modems without authentication, and have user privilege level 3. To improve device security, configure AUX login authentication.

The following are authentication modes available for modem dial-in through the console port:

- **None**—Requires no authentication and is insecure.
- **Password**—Requires a password for accessing the CLI. If your password was lost, log in to the device through the console port or modify the password.
- **Scheme**—Uses the AAA module to provide local or remote authentication. If your username or password was lost, log in to the device through the console port to modify the setting. If the username or password configured on a remote server was lost, contact the server administrator for help.

Table 16 Configuration required for different modem login authentication modes

Authentication mode	Configuration task	Reference
None	Set the authentication mode to none for the AUX user interface.	"Configuring none authentication for modem dial-in"
Password	Enable password authentication on the AUX user interface. Set a password.	"Configuring password authentication for modem dial-in"
Scheme	Enable scheme authentication on the AUX user interface. Configure local or remote authentication settings. To configure local authentication: <ol style="list-style-type: none"> 1. Configure a local user and specify the password. 2. Configure the device to use local authentication. To configure remote authentication: <ol style="list-style-type: none"> 3. Configure the RADIUS or HWTACACS scheme on the device. 4. Configure the username and password on the AAA server. 5. Configure the device to use the scheme for user authentication. 	"Configuring scheme authentication for modem dial-in"

Setting up the configuration environment

Set up a configuration environment as shown in Figure 19:

1. Connect the serial port of the PC to a modem and the console port of the device to a modem.
2. Connect each modem to the PSTN through a telephone cable.
3. Obtain the telephone number of the modem connected to the device.

Figure 19 Connecting the PC to the device through modems



4. Perform the following configurations on the modem directly connected to the device:
 - **AT&F**—Restores the factory default.
 - **ATS0=1**—Configures auto-answer on first ring.
 - **AT&D**—Ignores data Terminal Ready signals.
 - **AT&K0**—Disables local flow control.
 - **AT&R1**—Ignores Data Flow Control signals.
 - **AT&S0**—Forces **DSR** to remain on.
 - **ATEQ1&W**—Disables the modem from returning command responses and execution results.
 To verify your configuration, enter **AT&V** to display the configuration results.

NOTE:

The configuration commands and output vary by modem. For more information, see the modem user guide.

5. To avoid data loss, verify that the speed of the console port is lower than the transmission rate of the modem, and the default parity check, stop bits, and data bits settings are used.
6. Launch the terminal emulation program and create a connection by using the telephone number of the modem connected to the device.

Figure 20 to Figure 23 shows the configuration procedure in Windows XP HyperTerminal.

Figure 20 Creating a connection

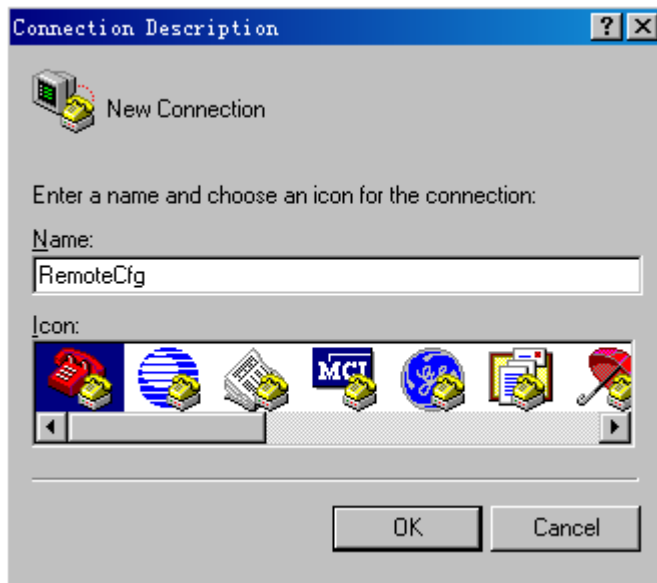
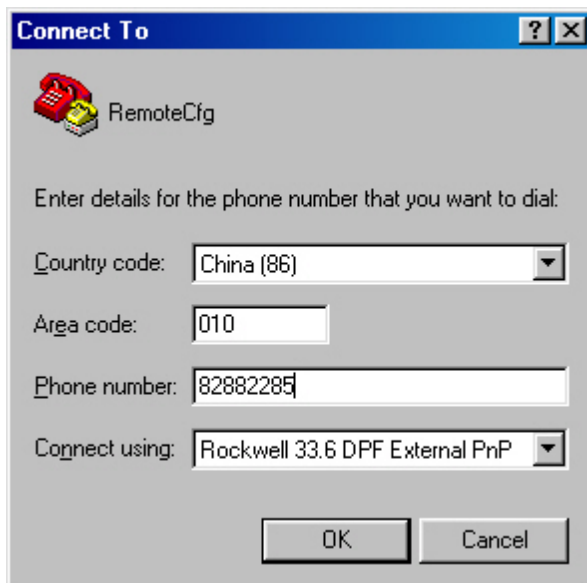


Figure 21 Configuring the dialing parameters

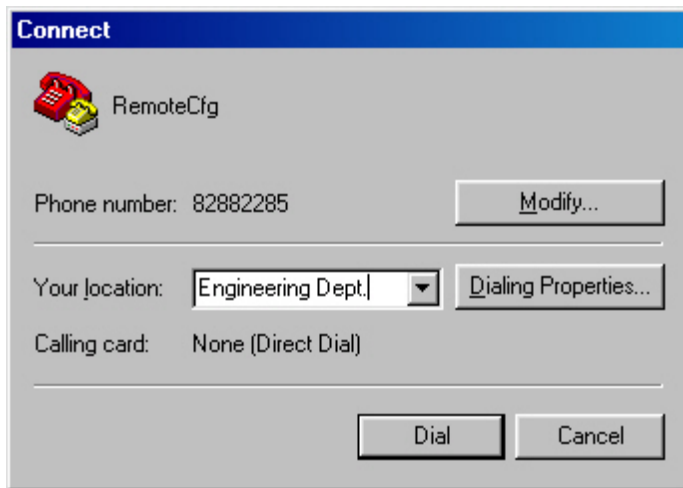


NOTE:

On Windows Server 2003, you must add the HyperTerminal program first, and then log in to and manage the device as described in this document. On Windows Server 2008, Windows 7, Windows Vista, or some other operating system, obtain a third-party terminal control program first, and follow the user guide or online help of that program to log in to the device.

7. Dial the telephone number to establish a connection to the device.

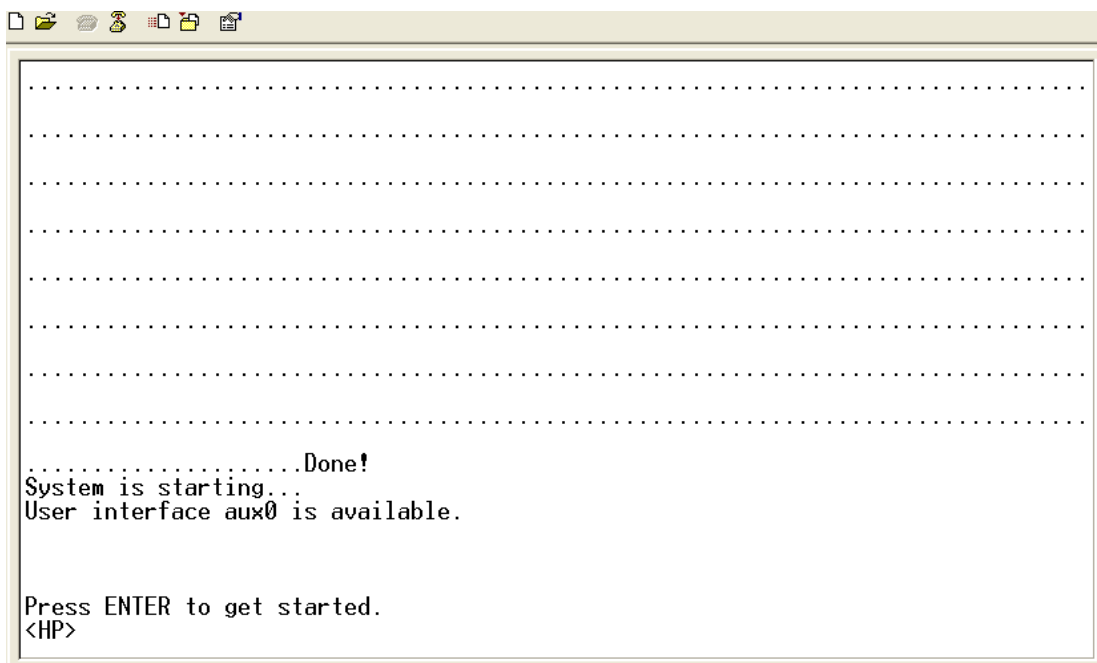
Figure 22 Dialing the number




Character string CONNECT9600 is displayed on the terminal.

8. Press **Enter** as prompted.

Figure 23 Configuration page



9. At the default user view prompt <HP>, enter commands to configure the device or view the running status of the device. To get help, enter ?.

To disconnect the PC from the device, execute the **ATH** command in the HyperTerminal. If the command cannot be entered, type AT+ + + and then press **Enter**. When the word "OK" appears, execute the **ATH** command. The connection is terminated if "OK" is displayed. You can also terminate the connection by clicking  in the HyperTerminal window.

! **IMPORTANT:**

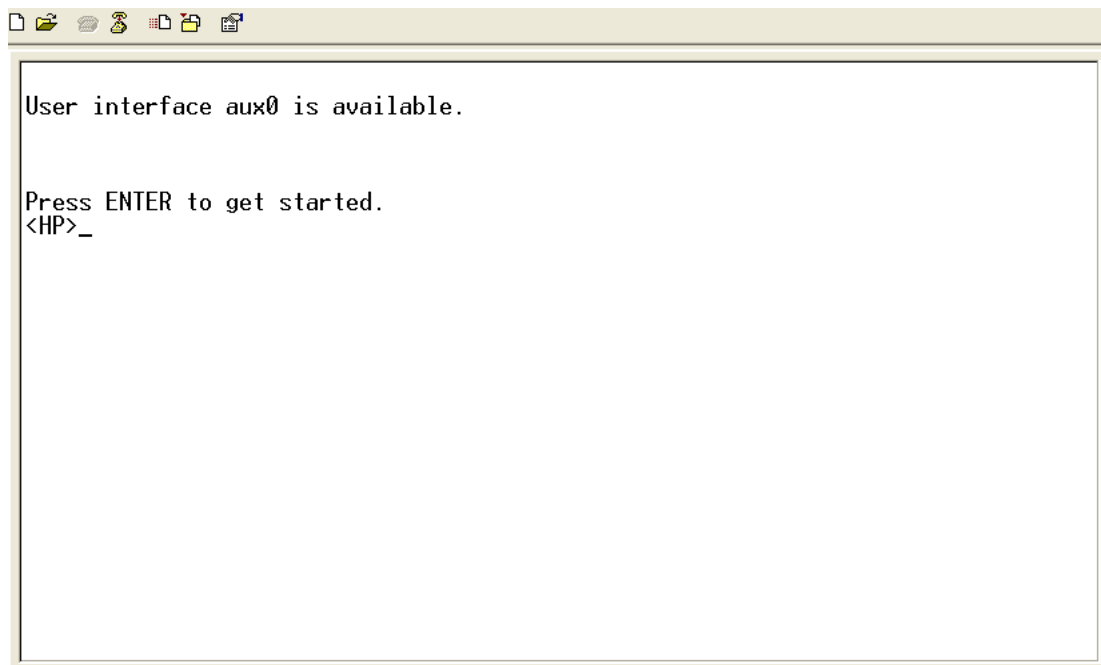
Do not directly close the HyperTerminal. Doing so can cause some modems to stay in use, and your subsequent dial-in attempts will always fail.

Configuring none authentication for modem dial-in

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter one or more AUX user interface views.	user-interface aux <i>first-number</i> [<i>last-number</i>]	N/A
3. Enable the none authentication mode.	authentication-mode none	By default, modem users can dial in to the device without authentication.
4. Configure common settings for the AUX user interfaces.	See " Configuring common settings for modem dial-in (optional) ."	Optional.

The next time you attempt to dial in to the device, you do not need to provide any username or password, as shown in [Figure 24](#).

Figure 24 Dialing in to the device without any authentication

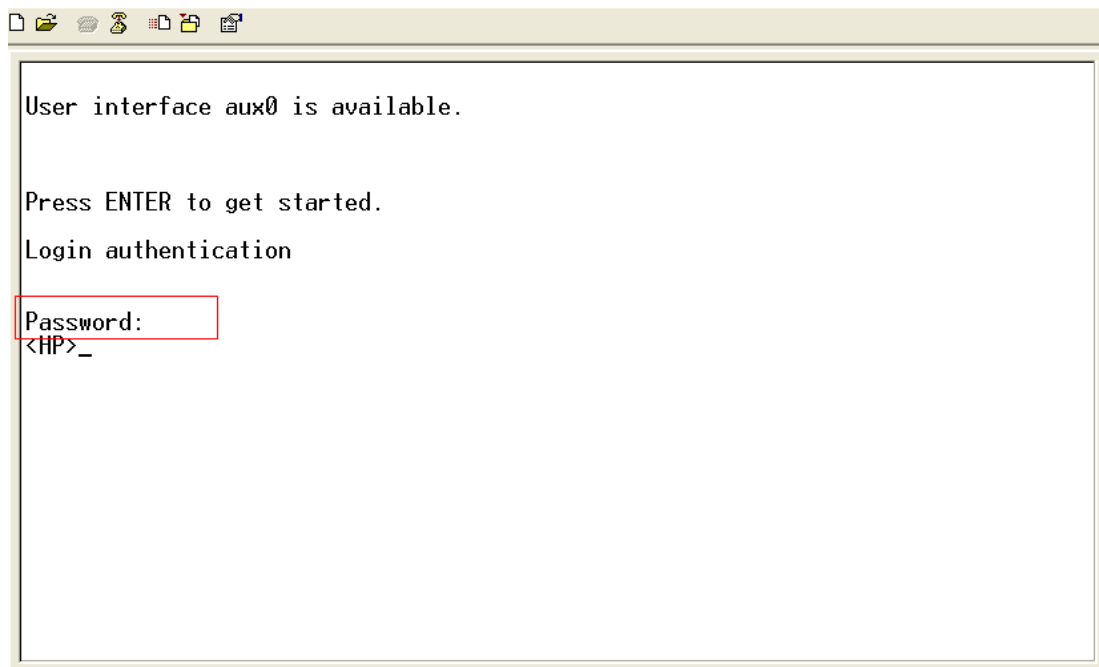


Configuring password authentication for modem dial-in

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter one or more AUX user interface views.	user-interface aux <i>first-number</i> [<i>last-number</i>]	N/A
3. Enable password authentication.	authentication-mode password	By default, no authentication is performed for modem dial-in users.
4. Set a password.	set authentication password { cipher simple } <i>password</i>	By default, no is set.
5. Configure common settings for the AUX user interfaces.	For more information, see " Configuring common settings for modem dial-in (optional) ."	Optional.

The next time you attempt to dial in to the device, you must provide the configured login password, as shown in [Figure 25](#).

Figure 25 Password authentication interface for modem dial-in users



Configuring scheme authentication for modem dial-in

Follow these guidelines when you configure scheme authentication for AUX login:

- To make the command authorization or command accounting function take effect, apply an HWTACACS scheme to the intended ISP domain. This scheme must specify the IP address of the authorization server and other authorization parameters.
- If the local authentication scheme is used, use the **authorization-attribute level** *level* command in local user view to set the user privilege level on the device.
- If a RADIUS or HWTACACS authentication scheme is used, set the user privilege level on the RADIUS or HWTACACS server.

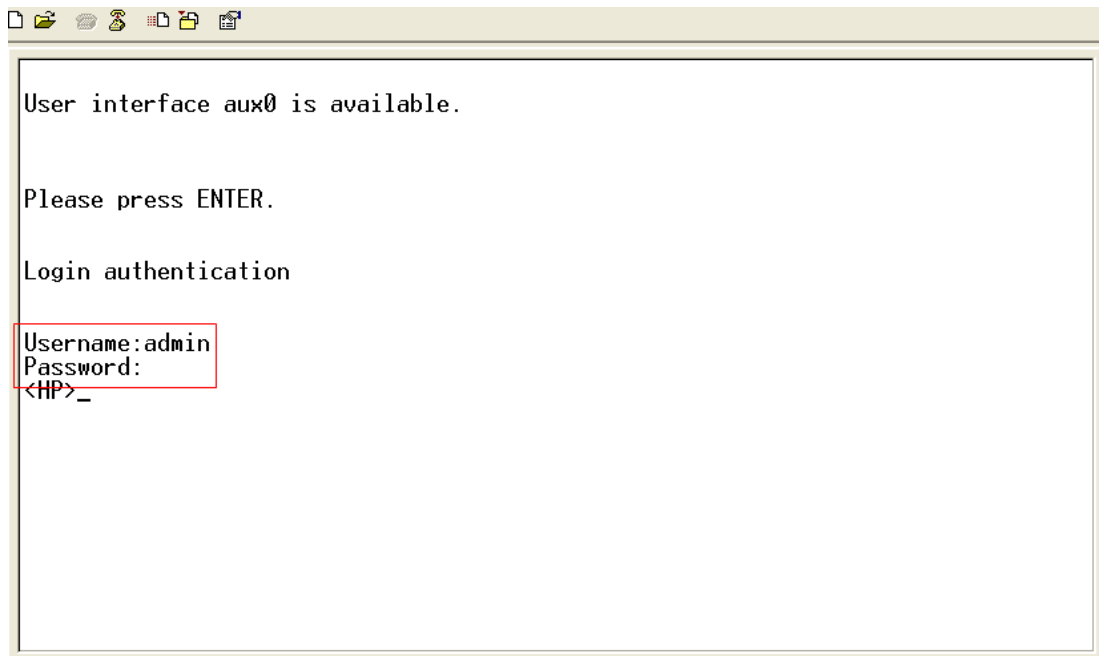
To configure scheme authentication for modem dial-in users:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter AUX user interface view.	user-interface aux <i>first-number</i> [<i>last-number</i>]	N/A
3. Enable scheme authentication.	authentication-mode scheme	Whether local, RADIUS, or HWTACACS authentication is adopted depends on the configured AAA scheme. By default, no authentication is performed for modem dial-in users.
4. Enable command authorization.	command authorization	Optional. By default, command authorization is disabled. The commands available for a user only depend on the user privilege level. If command authorization is enabled, a command is available only if the user has the commensurate user privilege level and is authorized to use the command by the AAA scheme.
5. Enable command accounting.	command accounting	Optional. By default, command accounting is disabled. The accounting server does not record the commands executed by users. Command accounting allows the HWTACACS server to record all executed commands that are supported by the device, regardless of the command execution result. This function helps control and monitor user behaviors on the device. If command accounting is enabled and command authorization is not enabled, every executed command is recorded on the HWTACACS server. If both command accounting and command authorization are enabled, only the authorized and executed commands are recorded on the HWTACACS server.
6. Exit to system view.	quit	N/A

Step	Command	Remarks
7. Apply an AAA authentication scheme to the intended domain.	<ol style="list-style-type: none"> 1. Enter the ISP domain view: domain <i>domain-name</i> 2. Apply the specified AAA scheme to the domain: authentication default { hwtacacs-scheme <i>hwtacacs-scheme-name</i> [local] local none radius-scheme <i>radius-scheme-name</i> [local] } 3. Exit to system view: quit 	<p>Optional.</p> <p>By default, local authentication is used.</p> <p>For local authentication, configure local user accounts.</p> <p>For RADIUS or HWTACACS authentication, configure the RADIUS or HWTACACS scheme on the device and configure authentication settings (including the username and password) on the server.</p> <p>For more information about AAA configuration, see <i>Security Configuration Guide</i>.</p>
8. Create a local user and enter local user view.	local-user <i>user-name</i>	By default, no local user exists.
9. Set a password for the local user.	password { cipher simple } <i>password</i>	By default, no password is set.
10. Specify the command level of the local user.	authorization-attribute level <i>level</i>	Optional. By default, the command level is 0.
11. Specify terminal service for the local user.	service-type terminal	By default, no service type is specified.
12. Configure common settings for the AUX user interfaces.	See " Configuring common settings for modem dial-in (optional) ."	Optional.

The next time you attempt to dial in to the device, you must provide the configured username and password, as shown in [Figure 26](#).

Figure 26 Scheme authentication interface for modem dial-in users



Configuring common settings for modem dial-in (optional)

⚠ CAUTION:

To avoid packet loss, make sure the speed of the console port is lower than the transmission rate of the modem.

Some common settings configured for an AUX user interface take effect immediately and can interrupt the login session. To save you the trouble of repeated re-logins, use a login method different from AUX login to log in to the device before you change AUX user interface settings.

After the configuration is complete, change the terminal settings on the configuration terminal and make sure they are the same as the settings on the device.

To configure common AUX user interface settings for modem dial-in accesses:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable copyright information display.	copyright-info enable	By default, copyright information display is enabled.
3. Enter one or more AUX user interface views.	user-interface aux <i>first-number</i> [<i>last-number</i>]	N/A
4. Configure the baud rate.	speed <i>speed-value</i>	By default, the baud rate is 9600 bps.
5. Configure the parity check mode.	parity { even none odd }	The default setting is none , namely, no parity check.

Step	Command	Remarks
6. Configure the number of stop bits.	stopbits { 1 1.5 2 }	The default is 1. Stop bits indicate the end of a character. The more the bits, the slower the transmission.
7. Configure the number of data bits in each character.	 databits { 7 8 }	By default, the number of data bits in each character is 8. The setting depends on the character coding type. For example, you can set it to 7 if standard ASCII characters are to be sent, and set it to 8 if extended ASCII characters are to be sent.
8. Define a shortcut key for starting a session.	activation-key <i>character</i>	By default, press Enter to start a session.
9. Define a shortcut key for terminating tasks.	escape-key { default <i>character</i> }	By default, press Ctrl+C to terminate a task.
10. Configure the flow control mode.	flow-control { hardware none software }	By default, the flow control mode is none . The device supports only the none mode.
11. Specify the terminal display type.	terminal type { ansi vt100 }	By default, the terminal display type is ANSI. The device supports two terminal display types: ANSI and VT100. HP recommends setting the display type to VT100 for both the device and the client. If the device and the client use different display types or both use the ANSI display type, when the total number of characters of a command line exceeds 80, the screen display on the terminal might be abnormal. For example, the cursor might be displayed at a wrong place.
12. Configure the user privilege level for login users.	user privilege level <i>level</i>	3 by default.
13. Set the maximum number of lines to be displayed on a screen.	screen-length <i>screen-length</i>	By default, a screen displays 24 lines at most. A value of 0 disables the function.
14. Set the size of the command history buffer.	history-command max-size <i>value</i>	By default, the buffer saves 10 history commands at most.

Step	Command	Remarks
15. Set the idle-timeout timer.	idle-timeout <i>minutes</i> [<i>seconds</i>]	The default idle-timeout is 10 minutes. The system automatically terminates the user's connection if there is no information interaction between the device and the user within the idle-timeout time. Setting idle-timeout to 0 disables the timer.

Displaying and maintaining CLI login

Task	Command	Remarks
Display information about the user interfaces that are being used.	display users [{ begin exclude include } <i>regular-expression</i>]	Available in any view.
Display information about all user interfaces the device supports.	display users all [{ begin exclude include } <i>regular-expression</i>]	Available in any view.
Display user interface information.	display user-interface [<i>num1</i> { aux vty } <i>num2</i>] [summary] [{ begin exclude include } <i>regular-expression</i>]	Available in any view.
Display the configuration of the device when it serves as a Telnet client.	display telnet client configuration [{ begin exclude include } <i>regular-expression</i>]	Available in any view.
Release a user interface.	free user-interface { <i>num1</i> { aux vty } <i>num2</i> }	Available in user view. Multiple users can log in to the system to simultaneously configure the device. You can execute the command to release the connections established on the specified user interfaces. You cannot use this command to release the connection you are using.
Lock the current user interface.	lock	Available in user view. By default, the current user interface is not locked.
Send messages to the specified user interfaces.	send { all <i>num1</i> { aux vty } <i>num2</i> }	Available in user view.

Logging in to the Web interface

The device provides a built-in Web server for you to configure the device through a Web browser. Web login is by default disabled.

To enable Web login, log in via the console port, and perform the following configuration tasks:

- Enable HTTP or HTTPS service.
- Configure the IP address of a Layer 3 interface, and make sure the interface and the configuration terminal can reach each other.
- Configure a local user account for Web login.

The device supports HTTP 1.0 and HTTPS for transferring webpage data across the Internet.

HTTPS uses SSL to encrypt data between the client and the server for data integrity and security, and is more secure than HTTP. You can define a certificate attribute-based access control policy to allow only legal clients to access the device.

HTTP login and HTTPS login are separate login methods. To use HTTPS login, you do not need to configure HTTP login.

Table 17 shows the basic Web login configuration requirements.

Table 17 Basic web login configuration requirements

Object	Requirements
Device	Configure an IP address for a Layer 3 interface. Configuring routes to make sure the interface and the PC can reach each other. Perform either or both of the following task: <ul style="list-style-type: none">• Configuring HTTP login• Configuring HTTPS login
PC	Install a Web browser. Obtain the IP address of the device's Layer 3 interface.

Configuring HTTP login

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable the HTTP service.	ip http enable	By default, HTTP service is enabled.
3. Configure the HTTP service port number.	ip http port <i>port-number</i>	Optional. The default HTTP service port is 80. If you execute the command multiple times, the last one takes effect.

Step	Command	Remarks
4. Associate the HTTP service with an ACL.	ip http acl <i>acl-number</i>	Optional. By default, the HTTP service is not associated with any ACL. Associating the HTTP service with an ACL enables the device to allow only clients permitted by the ACL to access the device.
5. Create a local user and enter local user view.	local-user <i>user-name</i>	By default, no local user is configured.
6. Configure a password for the local user.	password { cipher simple } <i>password</i>	By default, no password is configured for the local user.
7. Specify the command level of the local user.	authorization-attribute level <i>level</i>	No command level is configured for the local user.
8. Specify the Telnet service type for the local user.	service-type web	By default, no service type is configured for the local user.
9. Exit to system view.	quit	N/A
10. Set the DSCP value for IP to use for HTTP packets.	<ul style="list-style-type: none"> For IPv4: ip http dscp <i>dscp-value</i> For IPv6: ipv6 http dscp <i>dscp-value</i> 	Optional. The default is as follows: <ul style="list-style-type: none"> 16 for IPv4. 0 for IPv6.
11. Create a VLAN interface and enter its view.	interface vlan-interface <i>vlan-interface-id</i>	If the VLAN interface already exists, the command enters its view.
12. Assign an IP address and subnet mask to the interface.	ip address <i>ip-address</i> { <i>mask</i> <i>mask-length</i> }	By default, no IP address is assigned to the interface.

Configuring HTTPS login

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Associate the HTTPS service with an SSL server policy.	ip https ssl-server-policy <i>policy-name</i>	<p>By default, the HTTPS service is not associated with any SSL server policy, and the device uses a self-signed certificate for authentication.</p> <p>If you disable the HTTPS service, the system automatically de-associates the HTTPS service from the SSL service policy. Before re-enabling the HTTPS service, associate the HTTPS service with an SSL server policy first.</p> <p>If the HTTPS service has been enabled, any changes to the SSL server policy associated with the HTTP service that is enabled do not take effect.</p>

Step	Command	Remarks
3. Enable the HTTPS service.	ip https enable	<p>By default, HTTPS is disabled.</p> <p>Enabling the HTTPS service triggers an SSL handshake negotiation process. During the process, if the local certificate of the device exists, the SSL negotiation succeeds, and the HTTPS service can be started properly. If no local certificate exists, a certificate application process will be triggered by the SSL negotiation. Because the application process takes much time, the SSL negotiation often fails and the HTTPS service cannot be started normally. In that case, execute the ip https enable command multiple times to start the HTTPS service.</p>
4. Associate the HTTPS service with a certificate attribute-based access control policy.	ip https certificate access-control-policy <i>policy-name</i>	<p>Optional.</p> <p>By default, the HTTPS service is not associated with any certificate-based attribute access control policy.</p> <p>Associating the HTTPS service with a certificate-based attribute access control policy enables the device to control the access rights of clients.</p> <p>You must configure the client-verify enable command in the associated SSL server policy. If not, no clients can log in to the device.</p> <p>The associated SSL server policy must contain at least one permit rule. Otherwise, no clients can log in to the device.</p> <p>For more information about certificate attribute-based access control policies, see <i>Security Configuration Guide</i>.</p>
5. Specify the HTTPS service port number.	ip https port <i>port-number</i>	<p>Optional.</p> <p>The default HTTPS service port is 443.</p>
6. Associate the HTTPS service with an ACL.	ip https acl <i>acl-number</i>	<p>By default, the HTTPS service is not associated with any ACL.</p> <p>Associating the HTTPS service with an ACL enables the device to allow only clients permitted by the ACL to access the device.</p>
7. Create a local user and enter local user view.	local-user <i>user-name</i>	By default, no local user is configured.
8. Configure a password for the local user.	password { cipher simple } <i>password</i>	By default, no password is configured for the local user.
9. Specify the command level of the local user.	authorization-attribute level <i>level</i>	By default, no command level is configured for the local user.

Step	Command	Remarks
10. Specify the Web service type for the local user.	service-type web	By default, no service type is configured for the local user.
11. Exit to system view.	quit	N/A
12. Create a VLAN interface and enter its view.	interface vlan-interface <i>vlan-interface-id</i>	If the VLAN interface already exists, the command enters its view. You could replace this VLAN interface with any other Layer 3 interface as appropriate.
13. Assign an IP address and subnet mask to the interface.	ip address <i>ip-address</i> { <i>mask</i> <i>mask-length</i> }	By default, no IP address is assigned to the interface.

For more information about SSL and PKI, see *Security Configuration Guide*.

Displaying and maintaining Web login

Task	Command	Remarks
Display information about Web users.	display web users [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display HTTP state information.	display ip http [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display HTTPS state information.	display ip https [{ begin exclude include } <i>regular-expression</i>]	Available in any view

HTTP login configuration example

Network requirements

As shown in [Figure 27](#), configure the device to allow the PC to log in over the IP network by using HTTP.

Figure 27 Network diagram



Configuration procedure

1. Configure the device:


```
# Create VLAN 999, and add GigabitEthernet 1/0/1 (the interface connected to the PC) to VLAN 999.
<Sysname> system-view
```

```

[Sysname] vlan 999
[Sysname-vlan999] port GigabitEthernet 1/0/1
[Sysname-vlan999] quit

# Assign the IP address 192.168.0.58 and the subnet mask 255.255.255.0 to VLAN-interface 999.

[Sysname] interface vlan-interface 999
[Sysname-VLAN-interface999] ip address 192.168.0.58 255.255.255.0
[Sysname-VLAN-interface999] quit

# Create a local user named admin, and set the password to admin for the user. Specify the Web service type for the local user, and set the command level to 3 for this user.

[Sysname] local-user admin
[Sysname-luser-admin] service-type web
[Sysname-luser-admin] authorization-attribute level 3
[Sysname-luser-admin] password simple admin

```

2. Verify the configuration:

On the PC, run the Web browser. Enter the IP address of the device in the address bar. The Web login page appears, as shown in [Figure 28](#).

Figure 28 Web login page



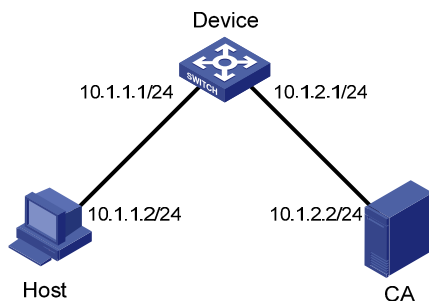
Enter the user name, password, verify code, select **English**, and click **Login**. The homepage appears. After login, you can configure device settings through the Web interface.

HTTPS login configuration example

Network requirements

As shown in [Figure 29](#), to prevent unauthorized users from accessing the device, configure the device as the HTTPS server and the host as the HTTPS client, and request a certificate for each of them.

Figure 29 Network diagram



Configuration procedure

This example assumes that the CA is named **new-ca**, runs Windows Server, and is installed with the Simple Certificate Enrollment Protocol (SCEP) add-on. This example also assumes that the device, host, and CA can reach one other.

1. Configure the device (HTTPS server):

Configure a PKI entity, configure the common name of the entity as **http-server1**, and the FQDN of the entity as **ssl.security.com**.

```
<Device> system-view
[Device] pki entity en
[Device-pki-entity-en] common-name http-server1
[Device-pki-entity-en] fqdn ssl.security.com
[Device-pki-entity-en] quit
```

Create a PKI domain, specify the trusted CA as **new-ca**, the URL of the server for certificate request as **http://10.1.2.2/certsrv/mscep/mscep.dll**, authority for certificate request as **RA**, and the entity for certificate request as **en**.

```
[Device] pki domain 1
[Device-pki-domain-1] ca identifier new-ca
[Device-pki-domain-1] certificate request url
http://10.1.2.2/certsrv/mscep/mscep.dll
[Device-pki-domain-1] certificate request from ra
[Device-pki-domain-1] certificate request entity en
[Device-pki-domain-1] quit
```

Create RSA local key pairs.

```
[Device] public-key local create rsa
```

Retrieve the CA certificate from the certificate issuing server.

```
[Device] pki retrieval-certificate ca domain 1
```

Request a local certificate from a CA through SCEP for the device.

```
[Device] pki request-certificate domain 1
```

Create an SSL server policy **myssl**, specify PKI domain 1 for the SSL server policy, and enable certificate-based SSL client authentication.

```
[Device] ssl server-policy myssl
[Device-ssl-server-policy-myssl] pki-domain 1
[Device-ssl-server-policy-myssl] client-verify enable
[Device-ssl-server-policy-myssl] quit
```

Create a certificate attribute group **mygroup1**, and configure a certificate attribute rule, specifying that the distinguished name (DN) in the subject name includes the string of **new-ca**.

```
[Device] pki certificate attribute-group mygroup1
[Device-pki-cert-attribute-group-mygroup1] attribute 1 issuer-name dn ctn new-ca
[Device-pki-cert-attribute-group-mygroup1] quit
```

Create a certificate attribute-based access control policy **myacp**. Configure a certificate attribute-based access control rule, specifying that a certificate is considered valid when it matches an attribute rule in certificate attribute group **myacp**.

```
[Device] pki certificate access-control-policy myacp
[Device-pki-cert-acp-myacp] rule 1 permit mygroup1
[Device-pki-cert-acp-myacp] quit
```

Associate the HTTPS service with SSL server policy **myssl**.

```
[Device] ip https ssl-server-policy myssl
```

Associate the HTTPS service with certificate attribute-based access control policy **myacp**.

```
[Device] ip https certificate access-control-policy myacp
```

Enable the HTTPS service.

```
[Device] ip https enable
```

Create a local user named **usera**, set the password to **123**, specify the Web service type, and specify the user privilege level 3. A level-3 user can perform all operations supported by the device.

```
[Device] local-user usera
[Device-luser-usera] password simple 123
[Device-luser-usera] service-type web
[Device-luser-usera] authorization-attribute level 3
```

2. Configure the host (HTTPS client):

On the host, run the IE browser, and then enter **http://10.1.2.2/certsrv** in the address bar and request a certificate for the host as prompted.

3. Verify the configuration:

Enter **https://10.1.1.1** in the address bar, and select the certificate issued by **new-ca**. When the Web login page of the device appears, enter the username **usera** and password **123** to log in to the Web management page.

For more information about PKI configuration commands, SSL configuration commands, and the **public-key local create rsa** command, see *Security Command Reference*.

Logging in through NMS

You can use an NMS to access the device MIB and perform GET and SET operations to manage and monitor the device. The device supports SNMPv1, SNMPv2c, and SNMPv3, and can work with various network management software products, including IMC. For more information about SNMP, see *Network Management and Monitoring Configuration Guide*.

By default, SNMP access is disabled. To enable SNMP access, log in to the device via any other method.

Configuring SNMP login

Connect the PC (the NMS) and the device to the network, making sure they can reach each other, as shown in Figure 30.

Figure 30 Network diagram



! IMPORTANT:

This document describes only the basic SNMP configuration procedures on the device. To make SNMP work correctly, make sure the SNMP settings (including the SNMP version) on the NMS are consistent with those on the device.

Prerequisites

- Assign an IP address to a Layer 3 interface on the device.
- Configure routes to make sure the NMS and the Layer 3 interface can reach each other.

Configuring SNMPv3 settings

Step	Command	Remarks
1. Enter system view.	<code>system-view</code>	N/A
2. Enable the SNMP agent.	<code>snmp-agent</code>	Optional. By default, the SNMP agent is disabled. You can enable SNMP agent with this command or any command that begins with <code>snmp-agent</code> .

Step	Command	Remarks
3. Configure an SNMP group and specify its access right.	snmp-agent group v3 <i>group-name</i> [authentication privacy] [read-view <i>read-view</i>] [write-view <i>write-view</i>] [notify-view <i>notify-view</i>] [acl <i>acl-number</i> acl ipv6 <i>ipv6-acl-number</i>] *	By default, no SNMP group is configured.
4. Add a user to the SNMP group.	snmp-agent usm-user v3 <i>user-name</i> <i>group-name</i> [[cipher] authentication-mode { md5 sha } <i>auth-password</i> [privacy-mode { 3des aes128 des56 } <i>priv-password</i>]] [acl <i>acl-number</i> acl ipv6 <i>ipv6-acl-number</i>] *	N/A

Configuring SNMPv1 or SNMPv2c settings

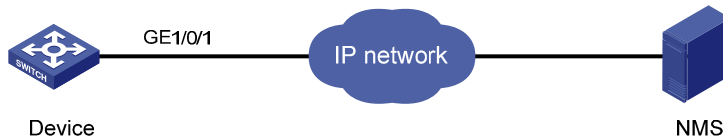
Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable the SNMP agent.	snmp-agent	Optional. By default, the SNMP agent is disabled. You can enable SNMP agent with this command or any command that begins with snmp-agent .
3. Create or update MIB view information.	snmp-agent mib-view { excluded included } <i>view-name</i> <i>oid-tree</i> [mask <i>mask-value</i>]	Optional. By default, the MIB view name is ViewDefault and OID is 1.
4. Configure SNMP NMS access right.	<ul style="list-style-type: none"> (Approach 1) Specify the SNMP NMS access right directly by configuring an SNMP community: snmp-agent community { read write } <i>community-name</i> [mib-view <i>view-name</i>] [acl <i>acl-number</i> acl ipv6 <i>ipv6-acl-number</i>] * (Approach 2) Configure an SNMP group and add a user to the SNMP group: <ul style="list-style-type: none"> a. snmp-agent group { v1 v2c } <i>group-name</i> [read-view <i>read-view</i>] [write-view <i>write-view</i>] [notify-view <i>notify-view</i>] [acl <i>acl-number</i> acl ipv6 <i>ipv6-acl-number</i>] * b. snmp-agent usm-user { v1 v2c } <i>user-name</i> <i>group-name</i> [acl <i>acl-number</i> acl ipv6 <i>ipv6-acl-number</i>] * 	Use either approach. The direct configuration approach is for SNMPv1 or SNMPv2c. The community name configured on the NMS should be consistent with the username configured on the agent. The indirect configuration approach is for SNMPv3.

NMS login example

Network requirements

Configure the device and network management station so you can remotely manage the device through SNMPv3.

Figure 31 Network diagram



Configuration procedure

1. Configure the device:

Assign an IP address to the device. Make sure the device and the NMS can reach each other. (Details not shown.)

Enter system view.

```
<Sysname> system-view
```

Enable the SNMP agent.

```
[Sysname] snmp-agent
```

Configure an SNMP group.

```
[Sysname] snmp-agent group v3 managev3group
```

Add a user to the SNMP group.

```
[Sysname] snmp-agent usm-user v3 managev3user managev3group
```

2. Configure the NMS:

Make sure the NMS has the same SNMP settings, including the username as the device. If not, the device cannot be discovered or managed by the NMS.

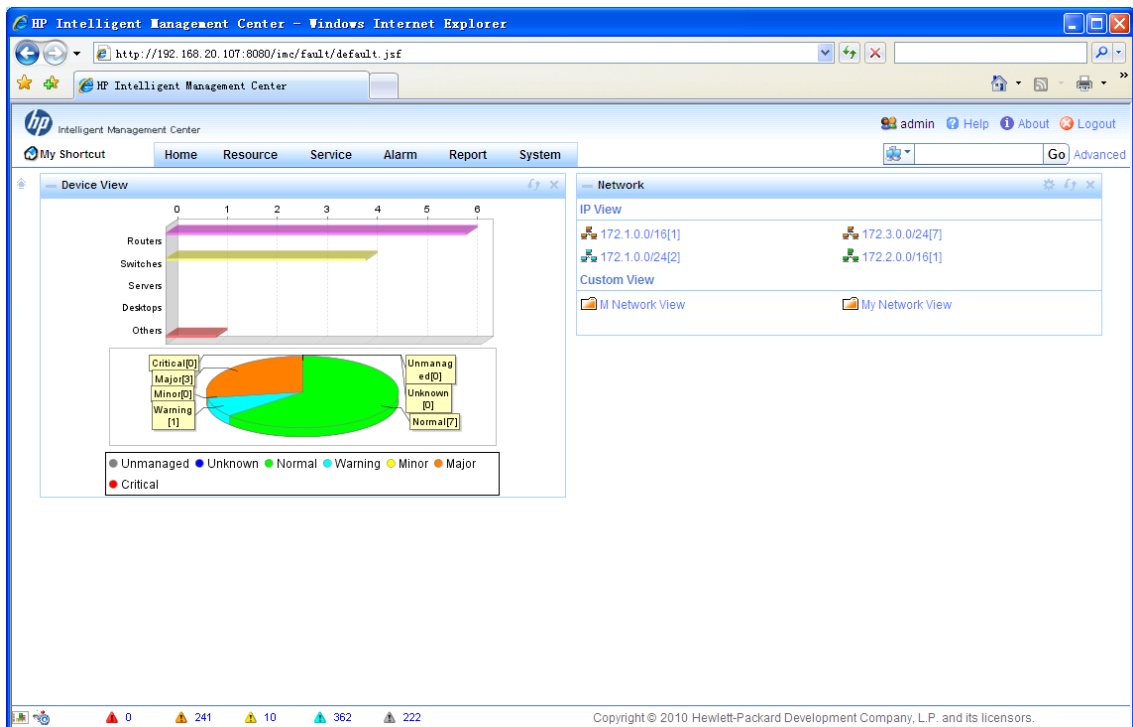
- a. On the PC, launch the browser, and enter **http://192.168.3.104:8080/imc** in the address bar (suppose that the IP address of IMC is 192.168.3.104).

Figure 32 IMC login page



- b. Enter the username and password, and then click **Login**.
The IMC homepage appears.

Figure 33 IMC homepage



- c. Log in to IMC and configure SNMP settings for IMC to find the switch.
- d. After the switch is found, you can manage and maintain the switch through IMC. For example, query switch information or configure switch parameters.

Controlling user logins

To harden device security, use ACLs to prevent unauthorized logins. For more information about ACLs, see *ACL and QoS Configuration Guide*.

Controlling Telnet logins

Use a basic ACL (2000 to 2999) to filter Telnet traffic by source IP address. Use an advanced ACL (3000 to 3999) to filter Telnet traffic by source and/or destination IP address. Use an Ethernet frame header ACL (4000 to 4999) to filter Telnet traffic by source MAC address.

To access the device, a Telnet user must match a permit statement in the ACL applied to the user interface.

Configuring source IP-based Telnet login control

Step	Command	Remarks
1. Enter system view.	<code>system-view</code>	N/A
2. Create a basic ACL and enter its view, or enter the view of an existing basic ACL.	<code>acl [ipv6] number acl-number [match-order { config auto }]</code>	By default, no basic ACL exists.
3. Configure an ACL rule.	<code>rule [rule-id] { permit deny } [source { sour-addr sour-wildcard any } time-range time-name fragment logging]*</code>	By default, a basic ACL does not contain any rule.
4. Exit the basic ACL view.	<code>quit</code>	N/A
5. Enter user interface view.	<code>user-interface [type] first-number [last-number]</code>	N/A
6. Use the ACL to control user logins by source IP address.	<code>acl [ipv6] acl-number { inbound outbound }</code>	<ul style="list-style-type: none">• inbound: Filters incoming packets.• outbound: Filters outgoing packets.

Configuring source/destination IP-based Telnet login control

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create an advanced ACL and enter its view, or enter the view of an existing advanced ACL.	acl [ipv6] number <i>acl-number</i> [match-order { config auto }]	By default, no advanced ACL exists.
3. Configure an ACL rule.	rule [<i>rule-id</i>] { permit deny } <i>rule-string</i>	N/A
4. Exit advanced ACL view.	quit	N/A
5. Enter user interface view.	user-interface [<i>type</i>] <i>first-number</i> [<i>last-number</i>]	N/A
6. Use the ACL to control user logins by source and destination IP addresses.	acl [ipv6] <i>acl-number</i> { inbound outbound }	<ul style="list-style-type: none"> • inbound: Filters incoming packets. • outbound: Filters outgoing packets.

Configuring source MAC-based Telnet login control

Ethernet frame header ACLs apply to Telnet traffic only if the Telnet client and server are located in the same subnet.

To configure source MAC-based Telnet login control:

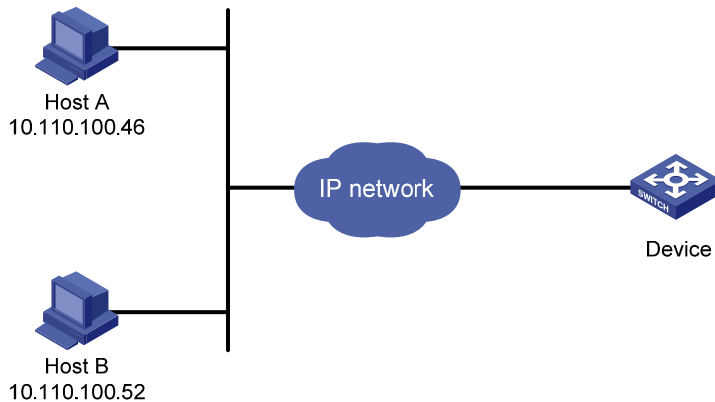
Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create an Ethernet frame header ACL and enter its view.	acl number <i>acl-number</i> [match-order { config auto }]	By default, no Ethernet frame header ACL exists.
3. Configure an ACL rule.	rule [<i>rule-id</i>] { permit deny } <i>rule-string</i>	N/A
4. Exit Ethernet frame header ACL view.	quit	N/A
5. Enter user interface view.	user-interface [<i>type</i>] <i>first-number</i> [<i>last-number</i>]	N/A
6. Use the ACL to control user logins by source MAC address.	acl <i>acl-number</i> inbound	inbound: Filters incoming packets.

Telnet login control configuration example

Network requirements

As shown in [Figure 34](#), configure an ACL on the device to permit only incoming Telnet packets sourced from Host A and Host B.

Figure 34 Network diagram



Configuration procedure

Configure basic ACL 2000, and configure rule 1 to permit packets sourced from Host B, and rule 2 to permit packets sourced from Host A.

```
<Sysname> system-view
[Sysname] acl number 2000 match-order config
[Sysname-acl-basic-2000] rule 1 permit source 10.110.100.52 0
[Sysname-acl-basic-2000] rule 2 permit source 10.110.100.46 0
[Sysname-acl-basic-2000] quit
```

Reference ACL 2000 in user interface view to allow Telnet users from Host A and Host B to access the Device.

```
[Sysname] user-interface vty 0 15
[Sysname-ui-vty0-15] acl 2000 inbound
```

Configuring source IP-based SNMP login control

Use a basic ACL (2000 to 2999) to control SNMP logins by source IP address. To access the requested MIB view, an NMS must use a source IP address permitted by the ACL.

Configuration procedure

To configure source IP-based SNMP login control:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a basic ACL and enter its view, or enter the view of an existing basic ACL.	acl [ipv6] number <i>acl-number</i> [match-order { config auto }]	By default, no basic ACL exists.
3. Create an ACL rule.	rule [<i>rule-id</i>] { permit deny } [source { <i>sour-addr</i> <i>sour-wildcard</i> any } time-range <i>time-name</i> fragment logging]*	N/A
4. Exit the basic ACL view.	quit	N/A

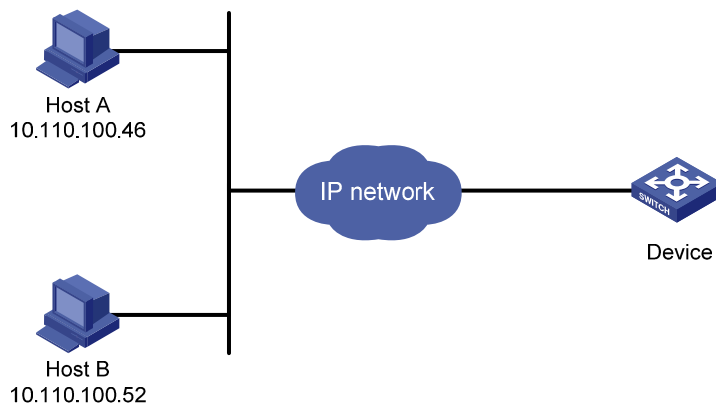
Step	Command	Remarks
5. Apply the ACL to an SNMP community, group or user.	<ul style="list-style-type: none"> SNMPv1/v2c community: snmp-agent community { read write } <i>community-name</i> [mib-view <i>view-name</i>] [acl <i>acl-number</i> acl ipv6 <i>ipv6-acl-number</i>] * SNMPv1/v2c group: snmp-agent group { v1 v2c } <i>group-name</i> [read-view <i>read-view</i>] [write-view <i>write-view</i>] [notify-view <i>notify-view</i>] [acl <i>acl-number</i> acl ipv6 <i>ipv6-acl-number</i>] * SNMPv3 group: snmp-agent group v3 <i>group-name</i> [authentication privacy] [read-view <i>read-view</i>] [write-view <i>write-view</i>] [notify-view <i>notify-view</i>] [acl <i>acl-number</i> acl ipv6 <i>ipv6-acl-number</i>] * SNMPv1/v2c user: snmp-agent usm-user { v1 v2c } <i>user-name</i> <i>group-name</i> [acl <i>acl-number</i> acl ipv6 <i>ipv6-acl-number</i>] * SNMPv3 user: snmp-agent usm-user v3 <i>user-name</i> <i>group-name</i> [[cipher] authentication-mode { md5 sha } <i>auth-password</i> [privacy-mode { 3des aes128 des56 } <i>priv-password</i>]] [acl <i>acl-number</i> acl ipv6 <i>ipv6-acl-number</i>] * 	For more information about SNMP, see <i>Network Management and Monitoring Configuration Guide</i> .

SNMP login control configuration example

Network requirements

As shown in Figure 35, configure the device to allow only NMS users from Host A and Host B to access.

Figure 35 Network diagram



Configuration procedure

Create ACL 2000, and configure rule 1 to permit packets sourced from Host B, and rule 2 to permit packets sourced from Host A.

```
<Sysname> system-view
[Sysname] acl number 2000 match-order config
[Sysname-acl-basic-2000] rule 1 permit source 10.110.100.52 0
[Sysname-acl-basic-2000] rule 2 permit source 10.110.100.46 0
[Sysname-acl-basic-2000] quit
```

Associate the ACL with the SNMP community and the SNMP group.

```
[Sysname] snmp-agent community read aaa acl 2000
[Sysname] snmp-agent group v2c groupa acl 2000
[Sysname] snmp-agent usm-user v2c usera groupa acl 2000
```

Configuring Web login control

Use a basic ACL (2000 to 2999) to filter HTTP traffic by source IP address for Web login control. To access the device, a Web user must use an IP address permitted by the ACL.

You can also log off suspicious Web users who have been logged in.

Configuring source IP-based Web login control

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a basic ACL and enter its view, or enter the view of an existing basic ACL.	acl [ipv6] number <i>acl-number</i> [match-order { config auto }]	By default, no basic ACL exists.
3. Create rules for this ACL.	rule [<i>rule-id</i>] { permit deny } [source { <i>sour-addr</i> <i>sour-wildcard</i> any } time-range <i>time-name</i> fragment logging]*	N/A
4. Exit the basic ACL view.	quit	N/A
5. Associate the HTTP service with the ACL.	ip http acl <i>acl-number</i>	N/A

Logging off online Web users

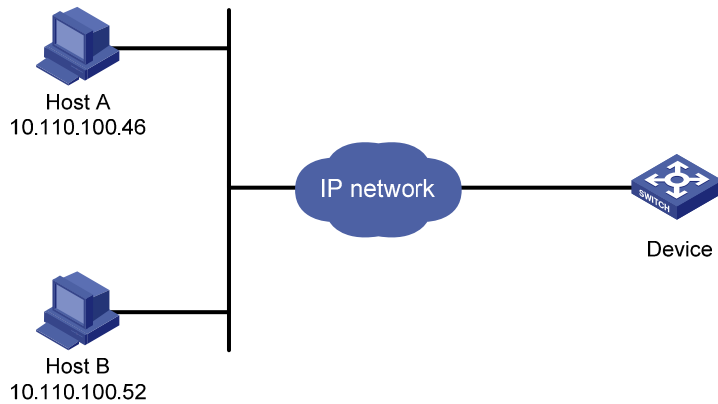
Task	Command	Remarks
Log off online Web users.	free web-users { all user-id <i>user-id</i> user-name <i>user-name</i> }	Available in user interface view

Web login control configuration example

Network requirements

As shown in [Figure 36](#), configure the device to allow only Web users from Host B to access.

Figure 36 Network diagram



Configuration procedure

Create ACL 2000, and configure rule 1 to permit packets sourced from Host B.

```
<Sysname> system-view
```

```
[Sysname] acl number 2030 match-order config
```

```
[Sysname-acl-basic-2030] rule 1 permit source 10.110.100.52 0
```

Associate the ACL with the HTTP service so only Web users from Host B are allowed to access the device.

```
[Sysname] ip http acl 2030
```

Configuring FTP

File Transfer Protocol (FTP) is an application layer protocol based on the client/server model. It is used to transfer files from one host to another over a TCP/IP network.

FTP server uses TCP port 20 to transfer data and TCP port 21 to transfer control commands. For more information about FTP, see RFC 959.

FTP supports the following transfer modes:

- **Binary mode**—Used to transfer image files, such as **.app** and **.bin** files.
- **ASCII mode**—Used to transfer text files, such as **.txt**, **.bat**, and **.cfg** files.

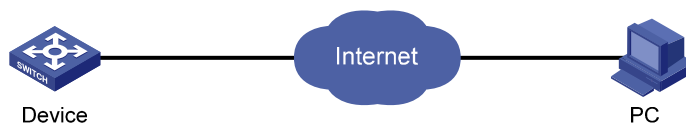
FTP can operate in either of the following modes:

- **Active mode (PORT)**—The FTP server initiates the TCP connection. This mode is not suitable when the FTP client is behind a firewall (for example, the FTP client resides in a private network).
- **Passive mode (PASV)**—The FTP client initiates the TCP connection. This mode is not suitable when the server does not allow the client to use a random unprivileged port greater than 1024.

The FTP operation mode varies depending on the FTP client program.

The device can act as the FTP client or FTP server:

Figure 37 FTP application scenario



Using the device as an FTP client

To connect to an FTP server or enter FTP client view, make sure the following requirements are met:

- You have level-3 (Manage) user privileges on the device. In FTP client view, whether a directory or file management command can be successfully executed depends on the authorization set on the FTP server.
- The device and the FTP server can reach each other.
- You have a user account (including the username, password, and authorization) on the FTP server. If the FTP server supports anonymous FTP, you can directly access the FTP server without a username and password.

Establishing an FTP connection

To access an FTP server, use the **ftp** command in user view or use the **open** command in FTP client view to establish a connection to the FTP server.

You can use the **ftp client source** command to specify a source IP address or source interface for the FTP packets sent by the device. If a source interface (typically a loopback interface) is specified, its primary IP address is used as the source IP address for the FTP packets sent by the device. The source interface setting and the source IP address setting overwrite each other.

The **ftp client source** command setting applies to all FTP sessions. When you set up an FTP session by using the **ftp** or **ftp ipv6** command, you can also specify a different source IP address for the FTP session.

! **IMPORTANT:**

To avoid FTP connection failures, when you specify a source interface for FTP packets, make sure that the interface has been assigned a primary IP address.

To establish an IPv4 FTP connection:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Specify a source IP address for outgoing FTP packets.	ftp client source { interface <i>interface-type interface-number</i> ip <i>source-ip-address</i> }	Optional. By default, the primary IP address of the output interface is used as the source IP address.
3. Return to user view.	quit	N/A
4. Log in to the FTP server.	<ul style="list-style-type: none"> • (Approach 1) Log in to the FTP server in user view: ftp [<i>server-address</i> [<i>service-port</i>] [vpn-instance <i>vpn-instance-name</i>] [source { interface <i>interface-type interface-number</i> ip <i>source-ip-address</i> }]] • (Approach 2) Log in to the FTP server in FTP client view: <ol style="list-style-type: none"> a. ftp b. open <i>server-address</i> [<i>service-port</i>] 	Use either approach. Only HP 5500-EI switches support the vpn-instance <i>vpn-instance-name</i> option.

To establish an IPv6 FTP connection, perform one of the following tasks:

Task	Command	Remarks
Log in to the FTP server from user view.	ftp ipv6 [<i>server-address</i> [<i>service-port</i>] [vpn-instance <i>vpn-instance-name</i>] [source ipv6 <i>source-ipv6-address</i>] [-i <i>interface-type interface-number</i>]]	Only HP 5500-EI switches support the vpn-instance <i>vpn-instance-name</i> option.
Log in to the FTP server from FTP client view.	<ol style="list-style-type: none"> 1. ftp ipv6 2. open ipv6 <i>server-address</i> [<i>service-port</i>] [-i <i>interface-type interface-number</i>] 	

Setting the DSCP value for IP to use for outgoing FTP packets

You can set the DSCP value for IPv4 or IPv6 to use for outgoing FTP packets on an FTP client, so that outgoing FTP packets are forwarded based on their priority on transit devices.

To set the DSCP value for IP to use for outgoing FTP packets:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the DSCP value for IP to use for outgoing FTP packets.	<ul style="list-style-type: none"> For IPv4: ftp client dscp <i>dscp-value</i> For IPv6: ftp client ipv6 dscp <i>dscp-value</i> 	The default is 0, whether the FTP client is running IPv4 or IPv6.

Managing directories on the FTP server

After the device establishes a connection to an FTP server, you can create or delete folders in the authorized directory on the FTP server.

To manage the directories on the FTP server:

Task	Command
Display detailed information about files and directories under the current directory on the FTP server.	dir [<i>remotefile</i> [<i>localfile</i>]]
Query a directory or file on the FTP server.	ls [<i>remotefile</i> [<i>localfile</i>]]
Change the working directory on the FTP server.	cd { <i>directory</i> .. / }
Return to the upper level directory on the FTP server.	cdup
Display the current directory on the FTP server.	pwd
Create a directory on the FTP server.	mkdir <i>directory</i>
Remove the specified working directory on the FTP server.	rmdir <i>directory</i>

Working with the files on the FTP server

After you log in to the server, you can upload a file to or download a file from the authorized directory by following these steps:

1. Use the **dir** or **ls** command to display the directory and the location of the file on the FTP server.
2. Delete unused files to get more free storage space.
3. Set the file transfer mode. FTP transmits files in two modes: ASCII and binary. Use ASCII mode to transfer text files. Use binary mode to transfer image files.
4. Use the **lcd** command to display the local working directory of the FTP client. You can upload the file or save the downloaded file in this directory.
5. Upload or download the file.

To work with the files on the FTP server:

Task	Command	Remarks
Display detailed information about a directory or file on the FTP server.	dir [<i>remotefile</i> [<i>localfile</i>]]	The ls command displays the name of a directory or file only, while the dir command displays detailed information such as the file size and creation time.

Task	Command	Remarks
Query a directory or file on the FTP server.	ls [<i>remotefile</i> [<i>localfile</i>]]	The ls command displays the name of a directory or file only, while the dir command displays detailed information such as the file size and creation time.
Delete the specified file on the FTP server permanently.	delete <i>remotefile</i>	N/A
Set the file transfer mode to ASCII.	ascii	By default, ASCII mode is used.
Set the file transfer mode to binary.	binary	By default, ASCII mode is used.
Set the FTP operation mode to passive.	passive	By default, passive mode is used
Display the local working directory of the FTP client.	lcd	N/A
Upload a file to the FTP server.	put <i>localfile</i> [<i>remotefile</i>]	N/A
Download a file from the FTP server.	get <i>remotefile</i> [<i>localfile</i>]	N/A

Switching to another user account

After you log in to the FTP server with one user account, you can switch to another user account to get a different privilege without reestablishing the FTP connection. You must correctly enter the new username and password. A wrong username or password can cause the FTP connection to disconnect.

To switch to another user account:

Task	Command
Change the username after FTP login.	user <i>username</i> [<i>password</i>]

Maintaining and troubleshooting the FTP connection

Task	Command	Remarks
Display the help information of FTP-related commands on the FTP server.	remotehelp [<i>protocol-command</i>]	N/A
Enable displaying detailed prompt information received from the server.	verbose	Enabled by default
Enable FTP related debugging when the device acts as the FTP client.	debugging	Disabled by default

Terminating the FTP connection

To terminate an FTP connection, perform one of the following tasks:

Task	Command	Remarks
Terminate the FTP connection without exiting FTP client view.	<ul style="list-style-type: none"> • disconnect • close 	Use either command in FTP client view.
Terminate the FTP connection and return to user view.	<ul style="list-style-type: none"> • bye • quit 	Use either command in FTP client view.

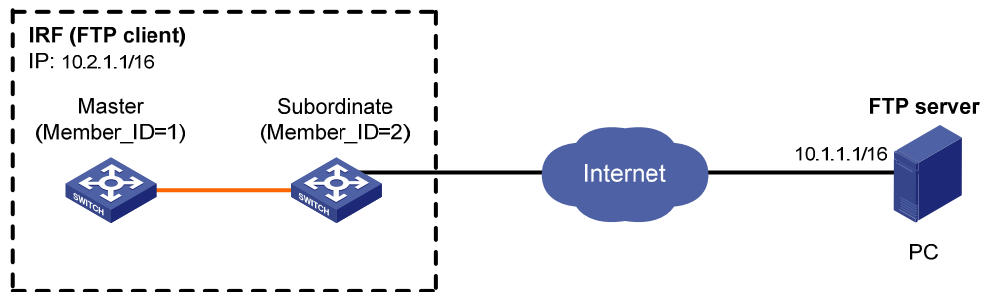
FTP client configuration example

Network requirements

As shown in Figure 38, the IRF fabric that comprises two member devices acts as the FTP client and the PC acts as the FTP server. The IRF fabric and the PC can reach each other. An account with the username **abc** and password **abc** is already configured on the FTP server.

Log in to the FTP server from the FTP client, download the file **newest.bin** from the FTP server to the FTP client, and upload the configuration file **config.cfg** from the FTP client to the FTP server for backup.

Figure 38 Network diagram



Note: The orange line represents an IRF link.

Configuration procedure

Examine the storage medium of the device for insufficiency or impairment. If no sufficient free space is available, use the **delete/unreserved file-url** command to delete unused files. (Details not shown.)

Log in to the server at 10.1.1.1 through FTP.

```
<Sysname> ftp 10.1.1.1
Trying 10.1.1.1 ...
Connected to 10.1.1.1.
220 WFTPD 2.0 service (by Texas Imperial Software) ready for new user
User(10.1.1.1:(none)):abc
331 Give me your password, please
Password:
230 Logged in successfully
```

Set the file transfer mode to binary.

```
[ftp] binary
200 Type set to I.
```

Download the system software image file **newest.bin** from the PC to the IRF fabric.

- Download the file **newest.bin** from the PC to the Flash root directory of the master device.

```
[ftp] get newest.bin
```

- Download the file **newest.bin** from the PC to the Flash root directory of the subordinate device (with member ID of 2).

```
[ftp] get newest.bin slot2#flash:/newest.bin
```

Set the transfer mode to ASCII and upload the configuration file **config.cfg** from the IRF fabric to the PC for backup.

```
[ftp] ascii
[ftp] put config.cfg back-config.cfg
227 Entering Passive Mode (10,1,1,1,4,2).
125 ASCII mode data connection already open, transfer starting for /config.cfg.
226 Transfer complete.
FTP: 3494 byte(s) sent in 5.646 second(s), 618.00 byte(s)/sec.
[ftp] bye
221 Server closing.
```

Specify **newest.bin** as the main system software image file for the next startup of all member devices.

```
<Sysname> boot-loader file newest.bin slot all main
This command will set the boot file of the specified board. Continue? [Y/N]:y
The specified file will be used as the main boot file at the next reboot on slot 1!
The specified file will be used as the main boot file at the next reboot on slot 2!
```

! **IMPORTANT:**

The system software image file used for the next startup must be saved in the Flash root directory. You can copy or move a file to the Flash root directory.

Reboot the device, and the system software image file is updated at the system reboot.

```
<Sysname> reboot
```

Using the device as an FTP server

If the device is operating as an FTP server, make sure the following requirements are met to ensure successful FTP operations:

- The device and the FTP server can reach each other.
- Configure a user account (including the username, password, and authorization) on the device or a remote authentication server for an FTP user. This task is required because the device does not support anonymous FTP for security reasons. By default, authenticated users can access the root directory of the device.
- The FTP user provides the correct username and password.

NOTE:

When you use the Internet Explorer browser to log in to the device operating as an FTP server, some FTP functions are not available. This is because multiple connections are required during the login process but the device supports only one connection at a time.

Configuring basic parameters

The FTP server uses one of the following modes to update a file when you upload the file (use the **put** command) to the FTP server:

- **Fast mode**—The FTP server starts writing data to the Flash after a file is transferred to the memory. This prevents the existing file on the FTP server from being corrupted in the event that anomaly, such as a power failure, occurs during a file transfer.
- **Normal mode**—The FTP server writes data to the Flash while receiving data. This means that any anomaly, such as a power failure, during file transfer might result in file corruption on the FTP server. This mode, however, consumes less memory space than the fast mode.

To configure basic parameters for the FTP server:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable the FTP server.	ftp server enable	By default, the FTP server is disabled.
3. Set the DSCP value for IPv4 to use for outgoing FTP packets.	ftp server dscp <i>dscp-value</i>	Optional. The default is 0.
4. Use an ACL to control FTP access.	ftp server acl <i>acl-number</i>	Optional. By default, no ACL is used for access control.
5. Configure the idle-timeout timer.	ftp timeout <i>minutes</i>	Optional. The default idle-timeout timer is 30 minutes. If no data is transferred within the idle-timeout time, the connection is terminated.
6. Set the file update mode for the FTP server.	ftp update { fast normal }	Optional. By default, normal update is used.
7. Return to user view.	quit	N/A
8. Release the FTP connection established by a specific user.	free ftp user <i>username</i>	Optional.

Configuring authentication and authorization

Perform this task on the FTP server to authenticate FTP clients and specify the directories that authenticated clients can access.

The following authentication modes are available:

- **Local authentication**—The device looks up the client's username and password in the local user account database. If a match is found, authentication succeeds.
- **Remote authentication**—The device sends the client's username and password to a remote authentication server for authentication. If this approach is used, the user account is configured on the remote authentication server rather than the device.

To assign an FTP user write access (including upload, delete, and create) to the device, assign level-3 (Manage) user privileges to the user. For read-only access to the file system, any user privilege level is OK.

For more information, see *Security Configuration Guide*.

To configure authentication and authorization for the FTP server:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a local user account and enter its view.	local-user <i>user-name</i>	By default, no local user account exists, and the system does not support FTP anonymous user access.
3. Set a password for the user account.	password { simple cipher } <i>password</i>	N/A
4. Assign FTP service to the user account	service-type ftp	By default, no service type is specified. If the FTP service is specified, the root directory of the device is by default used.
5. Configure authorization attributes.	authorization-attribute { acl <i>acl-number</i> callback-number <i>callback-number</i> idle-cut <i>minute</i> level <i>level</i> user-profile <i>profile-name</i> user-role { guest guest-manager security-audit } vlan <i>vlan-id</i> work-directory <i>directory-name</i> } *	Optional. By default, the FTP/SFTP users can access the root directory of the device, and the user level is 0. You can change the default configuration by using this command.

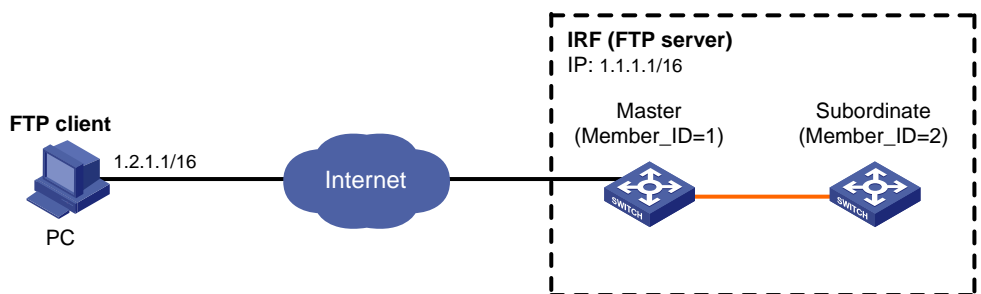
For more information about the **local-user**, **password**, **service-type ftp**, and **authorization-attribute** commands, see *Security Command Reference*.

FTP server configuration example

Network requirements

Create a local user account with username **abc** and password **abc** and enable FTP server on the IRF fabric in Figure 39. Use the user account to log in to the FTP server from the FTP client, upload the file **newest.bin** from the FTP client to the FTP server, and download the configuration file **config.cfg** from the FTP server to the FTP client for backup.

Figure 39 Network diagram



Note: The orange line represents an IRF link.

Configuration procedure

- Configure the FTP server:
 - # Examine the storage medium of the device for insufficiency or impairment. If no sufficient free space is available, use the **delete/unreserved file-url** command to delete unused files. (Details not shown.)

Create a local user account **abc**, set its password to **abc** and the user privilege level to level 3 (the manage level), specify the Flash root directory of the master device as the authorized directory, and specify the service type as FTP.

```
<Sysname> system-view
[Sysname] local-user abc
[Sysname-luser-abc] password simple abc
[Sysname-luser-abc] authorization-attribute level 3
[Sysname-luser-abc] authorization-attribute work-directory flash:/
[Sysname-luser-abc] service-type ftp
[Sysname-luser-abc] quit
```

To access the Flash root directory of the subordinate device (with the member ID 2), replace **flash:/** in the command **authorization-attribute work-directory flash:/** with **slot2#flash:/**.

Enable the FTP server.

```
[Sysname] ftp server enable
[Sysname] quit
```

2. Perform FTP operations from the FTP client:

Log in to the FTP server at 1.1.1.1 by using the username **abc** and password **abc**.

```
c:\> ftp 1.1.1.1
Connected to 1.1.1.1.
220 FTP service ready.
User(1.1.1.1:(none)):abc
331 Password required for abc.
Password:
230 User logged in.
```

Download the configuration file **config.cfg** from the FTP server to the PC for backup.

```
ftp> get config.cfg back-config.cfg
```

Upload the file **newest.bin** to the Flash root directory of the master.

```
ftp> put newest.bin
ftp> bye
```

This FTP procedure also applies to upgrading configuration files.

After you finish upgrading the Boot ROM image through FTP, execute the **bootrom update** command to upgrade Boot ROM.

3. Upgrade the FTP server:

Copy the system software image file **newest.bin** to the Flash root directory of the subordinate device (with the member ID 2).

```
<Sysname> copy newest.bin slot2#flash:/
```

Specify **newest.bin** as the main system software image file for the next startup of all member devices.

```
<Sysname> boot-loader file newest.bin slot all main
This command will set the boot file of the specified board. Continue? [Y/N]:y
The specified file will be used as the main boot file at the next reboot on slot
1!
The specified file will be used as the main boot file at the next reboot on slot
2!
```

! **IMPORTANT:**

The system software image file used for the next startup and the startup configuration file must be saved in the Flash root directory. You can copy or move a file to the Flash root directory.

```
# Reboot the IRF fabric and the system software image file is updated at the system reboot.  
<Sysname> reboot
```

Displaying and maintaining FTP

Task	Command	Remarks
Display the source IP address configuration of the FTP client.	display ftp client configuration [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the FTP server configuration.	display ftp-server [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display online FTP user information.	display ftp-user [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Configuring TFTP

Trivial File Transfer Protocol (TFTP) is a simplified version of FTP for file transfer over secure reliable networks. TFTP uses UDP port 69 for data transmission. In contrast to TCP-based FTP, TFTP requires no authentication or complex message exchanges, and is easier to deploy.

TFTP supports the following transfer modes:

- **Binary mode**—Used to transfer image files, such as **.app** and **.bin .btm** files.
- **ASCII mode**—Used to transfer text files, such as **.txt**, **.bat**, and **.cfg** files.

The device can operate only as a TFTP client (see Figure 40) to upload or download files.

Figure 40 TFTP application scenario



Prerequisites

Run a TFTP server program on the file host and set a TFTP working directory.

Configure IP addresses and routes to make sure that the device and the TFTP server can reach each other.

Using the device as a TFTP client

The device provides the following modes for downloading a new file from a TFTP server:

- **Normal download**—The new file is written directly to Flash and overwrites the old file that has the same name as it. If file download is interrupted, both old and new files are lost.
- **Secure download**—The new file is downloaded to memory and will not be written to Flash until the whole file is obtained. A download failure does not affect the old file that has the same name as the old file.

To avoid undesired file loss, use the secure download mode. If you use the normal download mode because of insufficient memory, assign the new file a file name unique in Flash.

You can use the **fttp client source** command to specify a source IP address or source interface for the TFTP packets sent by the device. If a source interface (typically, a loopback interface) is specified, its primary IP address is used as the source IP address for the TFTP packets. The source interface setting and the source IP address setting overwrite each other.

The **fttp client source** command setting applies to all TFTP sessions. When you set up a TFTP session with the **fttp** command, you can also specify a different source IP address for the TFTP session.

! IMPORTANT:

To avoid TFTP connection failures, when you specify a source interface for TFTP packets, make sure the interface has a primary IP address.

To configure the TFTP client:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Use an ACL to control the client's access to TFTP servers.	tftp-server [ipv6] acl acl-number	Optional. By default, no ACL is used for access control.
3. Specify a source IP address for outgoing TFTP packets.	tftp client source { interface interface-type interface-number ip source-ip-address }	Optional. By default, the primary IP address of the output interface is used as the source IP address.
4. Set the DSCP value for IP to use for outgoing TFTP packets.	<ul style="list-style-type: none"> For IPv4: tftp client dscp dscp-value For IPv6: tftp client ipv6 dscp dscp-value 	Optional The default is 0, whether the TFTP client is running IPv4 or IPv6.
5. Return to user view.	quit	N/A
6. Download or upload a file.	<ul style="list-style-type: none"> For IPv4: tftp server-address { get put sget } source-filename [destination-filename] [vpn-instance vpn-instance-name] [source { interface interface-type interface-number ip source-ip-address }] For IPv6: tftp ipv6 tftp-ipv6-server [-i interface-type interface-number] { get put } source-filename [destination-filename] [vpn-instance vpn-instance-name] 	Optional. Only HP 5500-EI switches support the vpn-instance vpn-instance-name option.

Displaying and maintaining the TFTP client

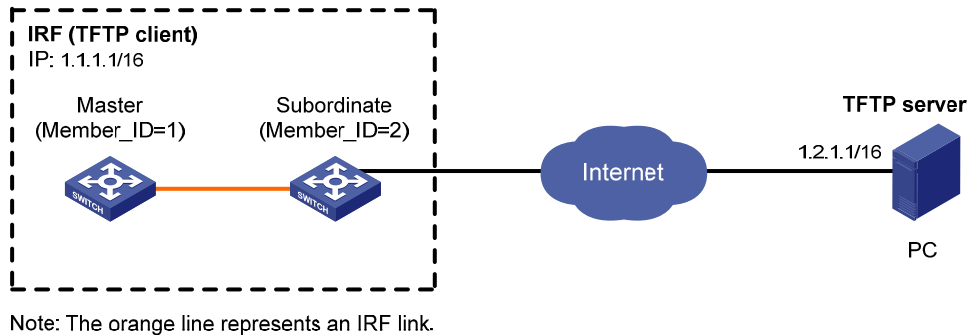
Task	Command	Remarks
Display the source IP address configuration of the TFTP client.	display tftp client configuration [{ begin exclude include } regular-expression]	Available in any view

TFTP client configuration example

Network requirements

Configure the PC in [Figure 41](#) as a TFTP server, and use TFTP to download the system software image file **newest.bin** from the TFTP server to the client and upload the configuration file **config.cfg** from the TFTP client to the server for backup.

Figure 41 Network diagram



Configuration procedure

This configuration procedure assumes that the PC and the IRF fabric can reach each other.

1. Configure the PC (TFTP server):

- Enable the TFTP server. (Details not shown.)
- Configure a TFTP working directory. (Details not shown.)

2. Configure the IRF fabric (TFTP client):

Examine the storage medium of the device for insufficiency or impairment. If no sufficient free space is available, use the **delete/unreserved file-url** command to delete unused files. (Details not shown.)

Download system software image file **newest.bin** from the PC to the master and subordinate devices.

- Download system software image file **newest.bin** from the PC to the root directory of the Flash on the master.

```
<Sysname> tftp 1.2.1.1 get newest.bin
```

- Download system software image file **newest.bin** from the PC to the root directory of the Flash on a subordinate device (with the member ID 2).

```
<Sysname> tftp 1.2.1.1 get newest.bin slot2#flash:/newest.bin
```

Upload a configuration file **config.cfg** to the TFTP server.

```
<Sysname> tftp 1.2.1.1 put config.cfg configback.cfg
```

Specify **newest.bin** as the main system software image file for the next startup for all member devices.

```
<Sysname> boot-loader file newest.bin slot all main
```

```
This command will set the boot file of the specified board. Continue? [Y/N]:y
```

```
The specified file will be used as the main boot file at the next reboot on slot 1!
```

```
The specified file will be used as the main boot file at the next reboot on slot 2!
```

! **IMPORTANT:**

The system software image file used for the next startup must be saved in the Flash root directory. You can copy or move a file to the Flash root directory.

Reboot the IRF fabric and the software is upgraded.

```
<Sysname> reboot
```

Managing the file system

This chapter describes how to manage the device's file system, including the storage media, directories and files.

Managing files

△ CAUTION:

To avoid file system corruption, do not plug or unplug storage media or perform active/standby switchover while the system is processing a file operation.

You can display directory or file information; display file contents; rename, copy, move, remove, restore, and delete files.

The copy operation enables you to create a file. You can also create a file by performing the download operation or using the **save** command.

When you specify a file, enter the file name in one of the formats shown in, Table 18.

Table 18 File name formats

Format	Description	Length	Example
<i>file-name</i>	Specifies a file in the current working directory.	1 to 91 characters	a.cfg indicates a file named a.cfg in the current working directory. This working directory might be on the master device or a subordinate device.
<i>path/file-name</i>	Specifies a file in a specific folder in the current working directory. The <i>path</i> argument represents the path to the file. If the file is in a single-level folder, specify the folder name for the argument. If the file is in a nested folder, separate each folder name by a forward slash (/).	1 to 135 characters	test/a.cfg indicates a file named a.cfg in the test folder in the current working directory.

Format	Description	Length	Example
<i>drive:/[path]/file-name</i>	<p>Specifies a file in a specific storage medium on the device.</p> <p>The <i>drive</i> argument represents the storage medium name.</p> <p>The storage medium on the master is typically flash.</p> <p>The storage medium on a subordinate device is typically slotX#flash, where X represents the member ID of the subordinate device, for example, slot2#flash.</p> <p>To view the correspondence between a device and its member ID, use the display irf command.</p>	1 to 135 characters	<p>flash:/test/a.cfg indicates a file named a.cfg in the test folder in the root directory of the Flash memory on the master.</p> <p>To access the file a.cfg in the root directory of the Flash on the subordinate device with member ID 2, enter slot2#flash:/a.cfg for the file name.</p>

Displaying file information

Perform this task in user view.

Task	Command
Display file or directory information.	dir [/all] [<i>file-url</i> /all-filestems]

Displaying file contents

Perform this task in user view.

Task	Command	Remarks
Display the contents of a file.	more <i>file-url</i>	Only text files can be displayed.

Renaming a file

Perform this task in user view.

Task	Command
Rename a file.	rename <i>fileurl-source fileurl-dest</i>

Copying a file

Perform this task in user view.

Task	Command
Copy a file.	copy <i>fileurl-source fileurl-dest</i>

Moving a file

Perform this task in user view.

Task	Command
Move a file.	move <i>fileurl-source fileurl-dest</i>

Deleting/restoring a file

You can delete a file permanently or just move it to the recycle bin. A file moved to the recycle bin can be restored, but a file permanently deleted cannot.

A file in the recycle bin occupies storage space. To release the occupied space, execute the **reset recycle-bin** command in the directory that holds the file. To save storage space, periodically empty the recycle bin with the **reset recycle-bin** command.

Perform the following tasks in user view:

Task	Command
Delete a file by moving it to the recycle bin.	delete <i>file-url</i>
Restore a file from the recycle bin.	undelete <i>file-url</i>
Delete a file permanently.	delete /unreserved <i>file-url</i>

Emptying the recycle bin

Step	Command	Remarks
1. Enter the original working directory of the file to be deleted in user view.	cd { <i>directory</i> <i>..</i> <i>/</i> }	Skip this step if the original directory of the file to be deleted is the current working directory.
2. Empty the recycle bin.	reset recycle-bin [<i>/force</i>]	N/A

Managing directories

You can create or remove a directory, display or change the current working directory, and display a specific directory.

Displaying directory information

Perform this task in user view.

Task	Command
Display directory or file information.	dir [<i>/all</i>] [<i>file-url</i> <i>/all-filestems</i>]

Displaying the current working directory

Perform this task in user view.

Task	Command
Display the current working directory.	<code>pwd</code>

Changing the current working directory

Perform this task in user view.

Task	Command
Change the current working directory.	<code>cd { <i>directory</i> .. / }</code>

Creating a directory

Perform this task in user view.

Task	Command
Create a directory.	<code>mkdir <i>directory</i></code>

Removing a directory

Before you remove a directory, you must delete all files and subdirectories in this directory. To delete a file, use the **delete** command; to delete a subdirectory, use the **rmdir** command.

The **rmdir** command automatically deletes the files in the recycle bin in the current directory.

Perform this task in user view.

Task	Command
Remove a directory.	<code>rmdir <i>directory</i></code>

Managing storage media

Storage media management includes space assignment. A storage medium is named based on the following rules:

If a storage medium is the only storage medium of its type on the device, it is named by its type. For example, if the device has only one Flash, the name of the Flash is **flash**.

Managing storage medium space

When the space of a storage medium becomes inaccessible, you can use the **fixdisk** command to examine the medium for damage and repair any damage.

The **format** command formats the storage medium, and all data on the storage medium is deleted.

△ CAUTION:

After a storage medium is formatted, all files on it are erased and cannot be restored. If a startup configuration file exists on the storage medium, formatting the storage medium results in loss of the startup configuration file.

To manage the space of a storage medium, perform the following tasks in user view:

Task	Command	Remarks
Repair a storage medium.	fixdisk <i>device</i>	N/A
Format a storage medium.	format <i>device</i>	N/A

Performing batch operations

A batch file comprises a set of executable commands. Executing a batch file is the same as executing the commands one by one. However, execution of a batch file does not guarantee successful execution of every command in the batch file. If a command has error settings or the conditions for executing the command are not satisfied, the system skips this command.

You can edit a batch file on your PC, and then upload or download it to the device. If the extension of the file is not **.bat**, use the **rename** command to change it to **.bat**.

To execute a batch file:

Step	Command
1. Enter system view.	system-view
2. Execute a batch file.	execute <i>filename</i>

Setting the file system operation mode

The file systems support the following operation modes:

- **alert**—The system warns you about operations that might cause problems such as file corruption and data loss. To prevent incorrect operations, use the **alert** mode.
- **quiet**—The system does not prompt for any operation confirmation.

To set the file system operation mode:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the file system operation mode.	file prompt { alert quiet }	Optional. The default is alert .

File system management examples

```
# Display the files and the subdirectories in the current directory.  
<Sysname> dir
```

```
Directory of flash:/
 0  -rw- 13308645 Mar 22 2011 11:34:07  main.bin
 1  -rw-   7380 Mar 25 2011 10:47:36  patch-package.bin
 2  -rw-   228 Mar 25 2011 10:50:39  patchstate
 3  -rw-   3921 Apr 01 2011 17:56:30  startup.cfg
 4  -rw-   151 Apr 01 2011 17:56:24  system.xml
15240 KB total (2521 KB free)

# Create new folder mytest in the test directory.
<Sysname> cd test
<Sysname> mkdir mytest
%Created dir flash:/test/mytest.

# Display the current working directory.
<Sysname> pwd
flash:/test

# Display the files and the subdirectories in the test directory.
<Sysname> dir
Directory of flash:/test/
 0  drw-      - Apr 01 2011 18:28:14  mytest
15240 KB total (2519 KB free)

# Return to the upper directory.
<Sysname> cd ..

# Display the current working directory.
<Sysname> pwd
flash:
```

Managing configuration files

You can manage configuration files at the CLI or by using the Boot menu of the device. This chapter describes the CLI approach.

Overview

A configuration file saves a configuration as a set of text commands. You can save the running configuration to a configuration file so the configuration takes effect after you reboot the device. You can also back up the configuration file to a host and download the file to the device as needed.

Configuration types

The device maintains the following types of configurations: factory defaults, startup configuration, and running configuration.

Factory defaults

The device is shipped with some basic settings called "factory defaults." These default settings make sure the device can start up and run normally when it has no configuration file or the configuration file is corrupted.

To view the factory defaults, use the **display default-configuration** command.

NOTE:

Factory defaults vary with device models and might differ from the default settings of commands.

Startup configuration

The device uses startup configuration to configure software features during startup. After the device starts up, you can specify a different configuration file for the next startup. This configuration file is called the "next-startup configuration file."

If no next-startup configuration file exists, the device boots with the factory defaults.

To view the current startup configuration, use either of the following ways:

- Execute the **display startup** command. To view detailed file contents, use the **more** command.
- After the device reboots, execute the **display current-configuration** command before making any configuration.

Running configuration

The running configuration includes the startup settings that have not been changed and the new settings you have made.

The running configuration is stored in a volatile storage medium. To make the settings you made to survive a reboot, save them to the startup configuration file.

To view the running configuration, use the **display current-configuration** command.

Configuration file format and content

A configuration file is saved as a text file according to the following rules:

- Commands are saved in their complete form.
- The commands are listed in sections by view, typically in this order: system view, interface view, protocol views, and user interface view.
- Sections are separated with one or more blank lines or comment lines that start with a pound sign (#).
- A configuration file ends with a return.

Coexistence of multiple configuration files

The device can save multiple configuration files on its storage media. You can save the configurations used for different networking environments to different configuration files. When the device moves between networking environments, you can quickly adapt the device to the environments by loading the intended configuration file onto the device.

You can specify one main startup configuration file and one backup startup configuration file for the device. At startup, the device first tries to start up with the main startup configuration file. If the main startup configuration file is corrupted or lost, the device tries to start up with the backup startup configuration file. For reliability, do not specify a configuration file as both the main and backup startup configuration files. If a configuration file is not assigned the main or backup attribute, its file attribute is NULL.

You can specify a main or backup startup configuration file directly (see "[Specifying a configuration file](#)") or when saving the running configuration (see "[Saving the running configuration](#)").

Startup with a configuration file

The device selects the configuration file to load at startup, as follows:

1. If the specified main startup configuration file exists, the device starts up with this configuration file.
2. If the specified main startup configuration file does not exist but the backup startup configuration file exists, the device starts up with the backup startup configuration file.
3. If neither the main nor the backup startup configuration file exists, the device starts up with the factory defaults.

Saving the running configuration

To make configuration changes take effect at the next startup of the device, save the running configuration to the startup configuration file to be used at the next startup before the device reboots.

Complete these tasks to save the current configuration:

Task	Remarks
Enabling configuration file auto-update	Optional Perform this task to ensure configuration consistency across member devices.
Saving running configuration in fast mode or safe mode	Required

Enabling configuration file auto-update

The configuration auto-update function enables all subordinate switches to automatically save the running configuration as the master does when you execute the **save** [**safely**] [**backup** | **main**] [**force**] command or the **save filename all** command. If this function is disabled, only the master saves the configuration.

To ensure configuration consistency, HP recommends enabling the function.

To enable configuration auto-update:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable configuration file auto-update.	slave auto-update config	By default, this function is enabled.

Saving running configuration in fast mode or safe mode

You can save the running configuration to a configuration file in one of the following modes:

- **Fast mode**—Use the **save** command without the **safely** keyword. The file is saved more quickly but is likely to be lost if the device reboots or power failure occurs during the process. If the startup configuration file for the next startup is lost, the device will use the factory defaults the next time it reboots, and you must re-specify a new startup configuration file for the device (see "Specifying a configuration file for the next startup").
- **Safe mode**—Use the **save** command with the **safely** keyword. The file is saved slowly, but the system retains the configuration file even if the device reboots or a power failure occurs during the process.

The fast saving mode is suitable for environments where a reliable power source is used. The safe mode is preferred in environments where the power source is not reliable or remote maintenance is involved.

The configuration file name extension must be **.cfg**.

To save the running configuration, perform either of the following tasks in any view:

Task	Command	Remarks
Save the running configuration to a configuration file without specifying the file as the startup configuration file for the next startup.	save file-url [all slot slot-number]	The save command executed with only the <i>file-url</i> argument saves the running configuration only to the specified path, regardless of whether the configuration auto-update function has been enabled.

Task	Command	Remarks
Save the running configuration to a configuration file and specify the file as the startup configuration file for the next startup.	save [safely] [backup main] [force]	By default, the running configuration is saved to the current startup configuration file. You can choose to save the configuration to a different file as instructed by the system. If you execute the save [safely] command without specifying any other keyword, the command saves the configuration to the main startup configuration file.

If configuration auto-update is enabled, the **save file-url all** command and the **save** [**safely**] [**backup** | **main**] [**force**] command save the configuration to the master device and all member devices. If the function is disabled, the commands save the configuration only to the master device.

Configuring configuration rollback

To replace the running configuration with the configuration in a configuration file without rebooting the device, use the configuration rollback function. This function helps you revert to a previous configuration state or adapt the running configuration to different network environments.

The configuration rollback function compares the running configuration against the specified replacement configuration file and handles configuration differences as follows:

- If a command in the running configuration is not in the replacement file, executes its **undo** form.
- If a command in the replacement file is not in the running configuration, adds it to the running configuration.
- If a command has different settings in the running configuration and the configuration file, replaces its running configuration with the setting in the configuration file.

To facilitate configuration rollback, the configuration archive function is developed. This function enables the system to automatically save the running configuration at regular intervals as checkpoint references.

Configuration task list

Task	Remarks
Configuring configuration archive parameters	Required.
<ul style="list-style-type: none"> • Enabling automatic configuration archiving • Manually archiving running configuration 	Required. Use either approach.
Performing configuration rollback	Required.

Configuring configuration archive parameters

Before archiving the running configuration, either manually or automatically, you must configure a file directory and file name prefix for configuration archives.

Configuration archives are saved with the file name format *prefix_serial number.cfg*, for example, **20080620archive_1.cfg** and **20080620archive_2.cfg**. The serial number is automatically assigned from 1 to 1000, increasing by 1. After the serial number reaches 1000, it restarts from 1.

After you change the file directory or file name prefix, or reboot the device, the old configuration archives are regarded as common configuration files, the configuration archive counter resets, and the **display archive configuration** command does not display them. The serial number for new configuration archives starts from 1.

After the maximum number of configuration archives is reached, the system deletes the oldest archive for the new archive.

Configuration guidelines

In an IRF fabric, the configuration archive function saves running configuration only on the master device. To make sure the system can archive running configuration after a master/subordinate switchover, create the directory on all IRF members.

Configuration procedure

To configure configuration archive parameters:

Step	Command	Remarks
1. Create the configuration archive directory.	See "Managing the file system."	In an IRF fabric, create the directory at least on the master. HP recommends creating the directory on all member devices.
2. Enter system view.	system-view	N/A
3. Configure the directory and file name prefix for archiving the running configuration.	archive configuration location <i>directory filename-prefix</i> <i>filename-prefix</i>	Do not include member ID information in the directory name. By default, no path or file name prefix is set for configuration archives, and the system does not regularly save configuration. ! IMPORTANT: The undo form of this command disables both manual and automatic configuration archiving, restores the default settings for the archive configuration interval and archive configuration max commands, and deletes all saved configuration archives.
4. Set the maximum number of configuration archives.	archive configuration max <i>file-number</i>	Optional. The default number is 10. Change the setting depending on the available storage space.

Enabling automatic configuration archiving

To avoid decreasing system performance, follow these guidelines when you configure automatic configuration archiving:

- If the device configuration does not change frequently, manually archive the running configuration as needed.
- If a low-speed storage media (such as a Flash) is used, archive the running configuration manually, or configure automatic archiving with an interval longer than 1440 minutes (24 hours).

Make sure you have set an archive path and file name prefix before performing this task.

To enable automatic configuration archiving:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable automatic configuration archiving and set the archiving interval.	archive configuration interval <i>minutes</i>	By default, this function is disabled. To view configuration archive names and their archiving time, use the display archive configuration command.

Manually archiving running configuration

To save system resources, disable automatic configuration archiving and manually archive configuration if the configuration will not be changed very often. You can also manually archive configuration before performing complicated configuration tasks so you can use the archive for configuration recovery after the configuration attempt fails.

Make sure you have set an archive path and file name prefix before performing this task.

Perform the following task in user view:

Task	Command
Manually archive the running configuration.	archive configuration

Performing configuration rollback

To avoid rollback failure, follow these guidelines:

- Do not reboot member devices while the system is executing the **configuration replace file** command. Make sure the replacement configuration file is created by using the configuration archive function or the **save** command on the current device.
- If the configuration file is not created on the current device, make sure the configuration file content format is fully compatible with the current device.
- The replacement configuration file is not encrypted.

To perform configuration rollback:

Step	Command
1. Enter system view.	system-view
2. Perform configuration rollback.	configuration replace file filename

The configuration rollback function might fail to reconfigure some commands in the running configuration for one of the following reasons:

- A command cannot be undone because the **undo** form designed for the command does not include a keyword or argument configured in the command. For example, if the **undo** form designed for the **A [B] C** command is **undo A C**, the configuration rollback function cannot undo the **A B C** command, because the system does not recognize the **undo A B C** command.
- A command (for example, a hardware-dependent command) cannot be deleted, overwritten, or undone due to system restrictions.
- The commands in different views are dependent on each other.
- Commands or command settings not supported on the current device cannot be added to the running configuration.

Specifying a configuration file for the next startup

You can specify a configuration file as the main startup configuration file to be used at the next startup when you use the **save** command to save the running configuration to it.

Alternatively, perform the following task in user view to specify a startup configuration file for the next startup:

Task	Command	Remarks
Specify a startup configuration file of all member switches.	startup saved-configuration <i>cfgfile</i> [backup main]	The configuration file must use the .cfg extension and be saved in the root directory of storage media.

Backing up the next-startup configuration file to a TFTP server

Before performing this task, make sure the server is reachable and enabled with TFTP service, and you have read and write permissions.

This task backs up only the main next-startup configuration file.

To back up the startup configuration file to be used at the next startup:

Step	Command	Remarks
1. Verify that a next-startup configuration file has been specified in user view.	display startup	Optional. If no next-startup configuration file has been specified, the back operation will fail.

Step	Command	Remarks
2. Back up the next-startup configuration file to a TFTP server in user view.	backup startup-configuration to <i>dest-addr</i> [<i>dest-filename</i>]	N/A

Deleting the next-startup configuration file

⚠ CAUTION:

This task permanently deletes the next-startup configuration file from all member devices. Before performing this task, back up the file as needed.

You can delete the main, the backup, or both. If the device has only one next-startup configuration file, the system sets the attribute of the configuration file to NULL instead of deleting the file.

You may need to delete the next-startup configuration file for one of the following reasons:

- After you upgrade system software, the file does not match the new system software.
- The file has been corrupted or is not fully compatible with the device.

After the file is deleted, the device uses factory defaults at the next startup.

Perform the following task in user view:

Task	Command
Delete the next-startup configuration file.	reset saved-configuration [backup main]

Restoring the next-startup configuration file from a TFTP server

To download a configuration file from a TFTP server to the root directory of each member' storage medium, and specify the file as the configuration file for the next startup, perform the task in this section.

This task restores only the main next-startup configuration file.

Before restoring the next-startup configuration file, make sure the server is reachable, the server is enabled with TFTP service, and you have read and write permissions.

To restore the next-startup configuration file from a TFTP server:

Step	Command	Remarks
1. Restore the main next-startup configuration file from a TFTP server in user view.	restore startup-configuration from <i>src-addr</i> <i>src-filename</i>	N/A
2. Verify that the specified configuration file has been set as the main next-startup configuration file.	display startup	Optional

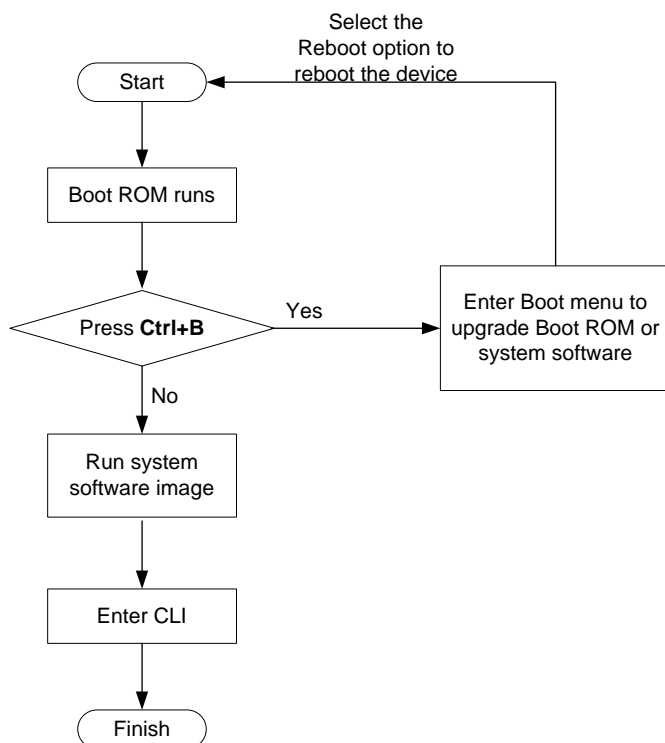
Displaying and maintaining a configuration file

Task	Command	Remarks
Display information about configuration rollback.	display archive configuration [{ begin exclude include } <i>regular-expression</i>]	Available in any view.
Display the running configuration.	display current-configuration [[configuration [<i>configuration</i>] interface [<i>interface-type</i>] [<i>interface-number</i>] exclude modules] [by-linenum] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view.
Display the factory defaults.	display default-configuration [{ begin exclude include } <i>regular-expression</i>]	Available in any view.
Display the running configuration file saved on the storage media of the device.	display saved-configuration [by-linenum] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view.
Display the configuration files used at this startup and the next startup.	display startup [[{ begin exclude include } <i>regular-expression</i>]	Available in any view.
Display the valid configuration under the current view.	display this [by-linenum] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view.

Upgrading software

Upgrading software includes upgrading the Boot ROM and system software. Each time the switch is powered on, it runs the Boot ROM image to initialize hardware and display hardware information, and then runs the system software image (also called the "boot file") so you can access the software features, as shown in Figure 42.

Figure 42 Relationship between the Boot ROM and the system software images



Software upgrade methods

You can upgrade both Boot ROM and system software at the Boot menu or at the CLI. The following sections describe the CLI approach. For instructions about how to upgrade them at the Boot menu, see the release notes of your switch.

The CLI approach provides the following upgrading methods:

Upgrade method	Upgrade object	Description
Upgrading software through a system reboot	Boot ROM and system software	A switch reboot is required during the upgrade, causing service interruption.
Upgrading software by installing hotfixes	System software	Hotfixes repair software defects without rebooting the switch and interrupting the running services of the switch. The patch files must match the switch model and software version. Otherwise, the hotfixing operation fails.

Upgrade method	Upgrade object	Description
In-Service Software Upgrade (ISSU)	System software	ISSU enables software upgrade and ensures continues packet forwarding. For more information, see "Performing ISSU." The ISSU feature is available only on the HP 5500-EI switch.

Upgrading software through a system reboot

Upgrading software by rebooting the switch interrupts the ongoing services. If any other method is possible, do not use this method.

Upgrading Boot ROM through a system reboot

1. Transfer the Boot ROM image to the root directory of the switch's storage media, for example, by using FTP or TFTP.

ⓘ IMPORTANT:

- To successfully upgrade the Boot ROM of a member switch, make sure the Boot ROM image is stored in the root directory of the member switch's storage media.

2. Upgrade the Boot ROM from the CLI, as follows:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable Boot ROM image validity check.	bootrom-update security-check enable	Optional By default, the validity check function is enabled. This feature examines the upgrade Boot ROM image for version and hardware incompatibility to ensure a successful upgrade.
3. Return to user view.	quit	N/A
4. Upgrade Boot ROM on member switches.	bootrom update file file-url slot slot-number-list	Available in user view.

3. Reboot the switch.

Upgrading system software through system reboot (method 1)

1. Transfer the system software image to the root directory of the master switch's storage media, for example, by using FTP or TFTP.

ⓘ IMPORTANT:

The image file must be saved in the root directory for a successful upgrade.

2. Copy the new system software image to the root directory of each subordinate switch's storage media.

❗ **IMPORTANT:**

You can assign different names to the image files for different member switches, but must make sure the image versions are the same.

3. Use the **boot-loader file** *file-url slot* { **all** | *slot-number* } { **main** | **backup** } command in user view to specify the system software image to be used at the next startup for each member switch.
4. Reboot all member switches.

Upgrading system software through system reboot (method 2)

This method simplifies the software upgrade procedure described in "Upgrading system software through system reboot (method 1)" for a multiple-MPU context by using one command to complete copying a system software image to an MPU and specifying the file as the system software image to be used at the next startup.

To use this method to upgrade system software:

1. Transfer the system software image to the root directory of the master switch's storage media, for example, by using FTP or TFTP.
2. Use the **boot-loader update file** *file-url slot* { *slot-number* | **all** } { **main** | **backup** } command in user view to specify the system software image as the file to be used at the next startup for each member switch.
3. Reboot all member switches.

Upgrading software by installing hotfixes

Hotfixes can repair software defects without rebooting the switch.

Basic concepts

- Patch and patch file

A patch fixes certain software defects. Patches might be released as patch files. A patch file might contain one or more patches. After being loaded from the storage media to the memory patch area, each patch is assigned a unique number, which starts from 1, for identification, management and operation. For example, if a patch file has three patches, they are numbered as 1, 2, and 3.
- Incremental patch

Incremental patches are dependent on previous patches and cannot separately run. For example, if a patch file has three patches, patch 3 can be running only after patch 1 and 2 take effect. You cannot run patch 3 separately.

Patches that have been released are all incremental patches.
- Common patch and temporary patch

There are common and temporary patches:

 - Common patches are formally released through the version release flow.
 - Temporary patches are not formally released through the version release flow, but temporarily provided to solve the emergent problems.

Common patches always include the functions of the previous temporary patches. The patch type only affects the patch loading process. The system deletes all of the temporary patches before it loads the common patch.

- Patch package file

A patch package file typically contains multiple patch files. A patch package file enables you to fix bugs for multiple components by executing a single command.

Using patch files is more troublesome. Patch file names are strictly defined and cannot be changed. To fix software bugs for a component, you must download the specific patch file, and rename the file to that pre-defined for the hardware. If the file name is unqualified, the upgrade fails. If there are multiple components, repeat this operation multiple times.

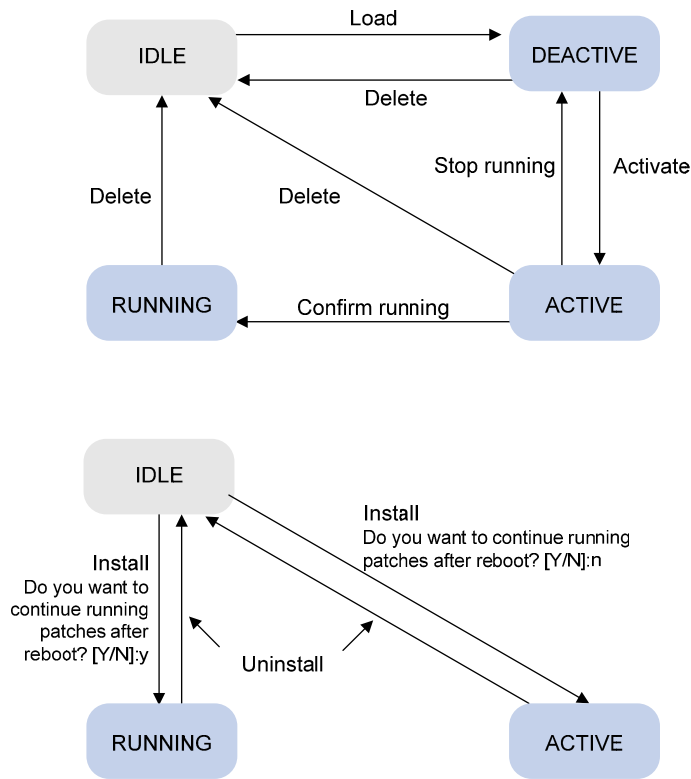
Patch state

Each patch has a state, which can be switched only by commands. The relationship between patch state changes and command actions is shown in Figure 43. The patch can be in IDLE, DEACTIVE, ACTIVE, or RUNNING state. Load, run temporarily, confirm running, stop running, delete, install, and uninstall are operations and they correspond to the following commands: **patch load**, **patch active**, **patch run**, **patch deactivate**, **patch delete**, **patch install**, and **undo patch install**. For example, if you execute the **patch active** command for the patches in DEACTIVE state, the patches switch to the ACTIVE state.

❗ **IMPORTANT:**

Patch state information is saved in Flash memory in the file **patchstate**. To make sure that the switch can correctly find the patches, do not edit, delete, move the file, or change the file name.

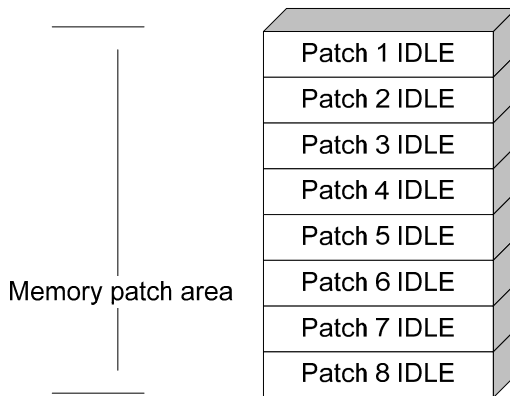
Figure 43 Relationship between patch state changes and command actions



IDLE state

Patches in IDLE state are not loaded. You cannot install or run the patches, as shown in [Figure 44](#). In this example, the memory patch area can load up to eight patches.

Figure 44 Patches are not loaded to the memory patch area

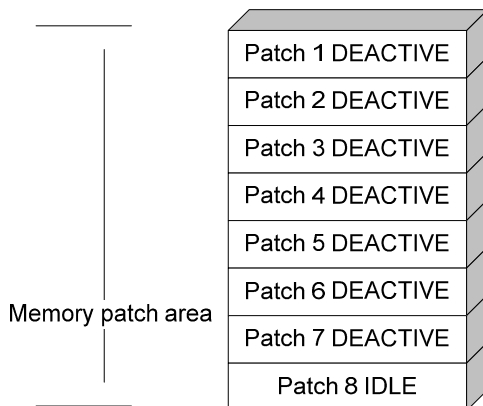


The memory patch area supports up to 200 patches.

DEACTIVE state

Patches in DEACTIVE state have been loaded to the memory patch area but have not yet run in the system. Suppose that there are seven patches in the patch file to be loaded. After the seven patches successfully pass the version check and CRC check, they are loaded to the memory patch area and are in DEACTIVE state. At this time, the patch states in the system are as shown in [Figure 45](#).

Figure 45 A patch file is loaded to the memory patch area

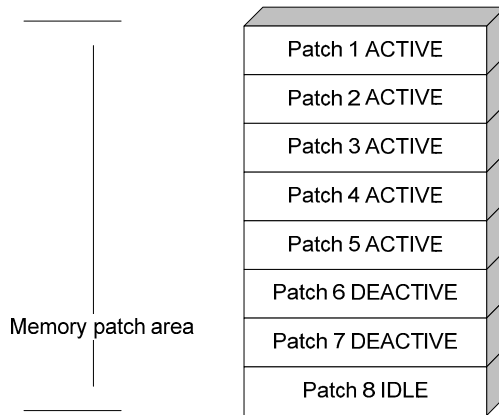


ACTIVE state

Patches in ACTIVE state have run temporarily in the system and become DEACTIVE after system reboot. For the seven patches in [Figure 45](#), if you activate the first five patches, their states change from DEACTIVE to ACTIVE. The patch states in the system are as shown in [Figure 46](#).

The patches that are in ACTIVE state are in DEACTIVE state after system reboot.

Figure 46 Patches are activated

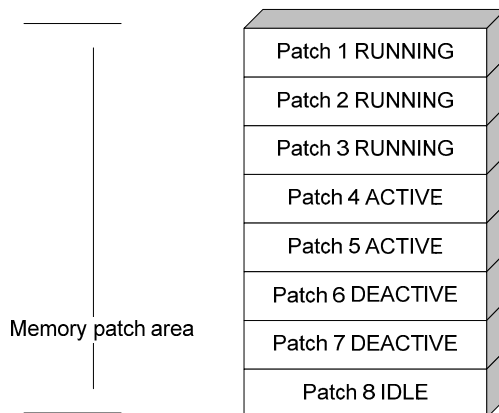


RUNNING state

After you confirm the ACTIVE patches are running, the state of the patches changes to RUNNING and the patches are in RUNNING state after system reboot. If you confirm the first three patches are running, for the five patches in [Figure 46](#), their states change from ACTIVE to RUNNING. The patch states of the system are shown in [Figure 47](#).

The patches that are in RUNNING state are still in RUNNING state after system reboot.

Figure 47 Patches are running



Hotfix configuration task list

Task	Remarks
Installing patches: <ul style="list-style-type: none">• Installing a patch in one step• Installing a patch step-by-step	Use either approach. The step-by-step patch installation allows you to control the patch status.
Uninstalling a patch step-by-step	Optional.

Installation prerequisites

Patches are released per switch model. To ensure a successful hotfix operation and normal switch operation after the hotfix operation:

- Make sure each patch file you are installing matches the switch model and software version.
- The loading and installation are performed on all member switches of an IRF fabric. Before these operations, save the same patch files to the storage media's root directory of each member switch.
- Name a patch file properly. Otherwise, the system cannot locate the patch file and the hotfixing operation fails. The name is in the format of "patch_PATCH-FLAG suffix.bin". The PATCH-FLAG is pre-defined. The value of the version field (using the **display patch information** command) represents the PATCH-FLAG suffix. The system searches the root directory of the storage media (Flash by default) for patch files based on the PATCH-FLAG. If there is a match, the system loads patches to or install them on the memory patch area. The default name of a patch file is patch_xxx.bin.

Installing a patch in one step

To install patches in one step, use the **patch install** command and specify either the directory where the patch file locates or the filename of the patch package.

After you execute the command, the system displays the message "Do you want to continue running patches after reboot? [Y/N]:"

- Entering **y** or **Y**: All the specified patches are installed, and turn to RUNNING state from IDLE. This equals executing commands **patch location**, **patch load**, **patch active**, and **patch run**. The patches remain RUNNING after system reboot.
- Entering **n** or **N**: All the specified patches are installed and turn to ACTIVE state from IDLE. This equals executing commands **patch location**, **patch load** and **patch active**. The patches turn to DEACTIVE state after system reboot.

To install the patches in one step:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Install the patches in one step.	patch install { <i>patch-location</i> file <i>patch-package</i> }	<ul style="list-style-type: none">• <i>patch-location</i>: Specifies the directory where the patch file locates. Provide this argument when you install a patch file which is not packaged in a patch package file.• file <i>patch-location</i>: Specifies the name of the patch package file. Provide this option when you install a patch package file.

If you execute the **patch install** *patch-location* command, the directory specified for the *patch-location* argument will replace the directory specified with the **patch location** command after the upgrade is complete.

If you execute the **patch install file** *patch-package* command, the directory specified with the **patch location** command does not change.

To uninstall all patches in one operation, use the **undo patch install** command, which is the same as performing [Uninstalling a patch step-by-step](#).

Installing a patch step-by-step

Step-by-step patch installation enables you to control the patch status during the patch installation process.

Step-by-step patch installation task list

Task	Remarks
Configuring the patch file location	Optional. To install a patch package, skip this step.
Loading a patch file	Required.
Activating patches	Required.
Confirming running patches	Optional.

Configuring the patch file location

HP recommends that you save the patch file to the root directory of the Flash. If you save the patch files to a storage medium other than the Flash, you must specify the directory that saves patch files so the system can locate them.

Make sure the specified directory exist on each member switch in the IRF fabric.

If the switch has only one storage medium, you do not need to perform this task.

To configure the patch file location:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the patch file location.	patch location <i>patch-location</i>	Optional. flash: by default.

NOTE:

After you execute the **patch install** *patch-location* command, the directory specified for the *patch-location* argument will replace the directory specified with the **patch location** command after the upgrade is complete.

Loading a patch file

Loading the correct patch files is the basis of other hotfix operations.

If you install a patch from a patch file, the system loads a patch file from the Flash by default.

If you install a patch from a patch package, the system finds the correct patch file in the patch package file and loads the patch file.



IMPORTANT:

Set the file transfer mode to binary mode before using FTP or TFTP to upload or download patch files to or from the Flash of the switch. Otherwise, patch file cannot be parsed properly.

To load a patch file:

Step	Command
1. Enter system view.	system-view
2. Load the patch file on from the storage media (the Flash) to the specified memory patch area.	patch load slot <i>slot-number</i> [file <i>patch-package</i>]

Activating patches

After you activate a patch, the patch takes effect and is in the test-run stage. After the switch is reset or rebooted, the patch becomes invalid.

If you find that an ACTIVE patch causes an error, reboot the switch to deactivate the patch, so as to avoid a series of running faults resulting from patch error.

To activate patches:

Step	Command
1. Enter system view.	system-view
2. Activate patches.	patch active [<i>patch-number</i>] slot <i>slot-number</i>

Confirming running patches

This operation is applicable only to patches in ACTIVE state.

After you confirm that the installed patch is running, the patch state changes to RUNNING, and the patch is in the normal running stage. After the switch is reset or rebooted, the patch is still valid.

To confirm the running of patches:

Step	Command
1. Enter system view.	system-view
2. Confirm the running of patches.	patch run [<i>patch-number</i>] [slot <i>slot-number</i>]

Uninstalling a patch step-by-step

This section describes the procedure of uninstalling patches.

Step-by-step patch uninstallation task list

Task	Remarks
Stopping running patches	Required
Deleting patches	Required

Stopping running patches

When you stop running a patch, the patch state becomes DEACTIVE, and the system runs the way it did before it was installed with the patch.

To stop running patches:

Step	Command
1. Enter system view.	system-view
2. Stop running patches.	patch deactivate [<i>patch-number</i>] slot <i>slot-number</i>

Deleting patches

Deleting patches only removes the patches from the memory patch area, and does not delete them from the storage media. The patches turn to IDLE state after this operation. After a patch is deleted, the system runs the way it did before it was installed with the patch.

In an IRF fabric, HP recommends that you uninstall all patches by using the **undo patch install** command in one operation.

To delete patches:

Step	Command
1. Enter system view.	system-view
2. Delete patches from the memory patch area.	patch delete [<i>patch-number</i>] slot <i>slot-number</i>

Displaying and maintaining software upgrade

Task	Command	Remarks
Display information about system software	display boot-loader [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about the patch package.	display patch [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the patch information.	display patch information [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Software upgrade examples

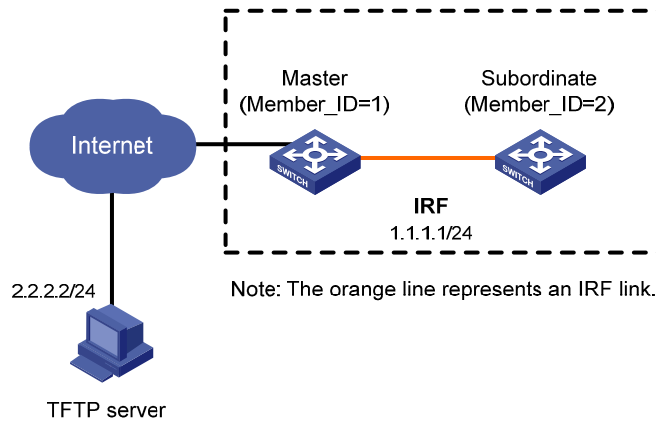
Immediate upgrade configuration example

Network requirement

The IRF fabric in [Figure 48](#) comprises two member switches, the master use the member ID 1 and the subordinate switch uses the member ID 2. The current software version of the IRF fabric is **soft-version1**. The latest system software image **soft-version2.bin** and the latest configuration file **new-config.cfg** are both saved on the TFTP server. The TFTP server and IRF fabric can reach each other.

Upgrade the software version of the IRF fabric to **soft-version2** and the configuration file to **new-config**.

Figure 48 Network diagram



Configuration procedure

1. Configure the TFTP server (the configuration varies with server vendors):
Obtain the system software image and configuration file through legitimate channels, such as the official website of HP, agents, and technical staff. Save these files under the TFTP server's working path for the access of the TFTP clients.
2. Configure the members of the IRF fabric:

Download file **new-config.cfg** on the TFTP server to the master.

```
<IRF> tftp 2.2.2.2 get new-config.cfg
..
File will be transferred in binary mode
Downloading file from remote TFTP server, please wait.....
TFTP: 917 bytes received in 1 second(s)
File downloaded successfully.
```

Download file **new-config.cfg** to the subordinate switch with the member ID of 2.

```
<IRF> tftp 2.2.2.2 get new-config.cfg slot2#flash:/new-config.cfg
```

Download file **soft-version2.bin** on the TFTP server to the master and subordinate switch.

```
<IRF> tftp 2.2.2.2 get soft-version2.bin
...
File will be transferred in binary mode
Downloading file from remote TFTP server, please wait.....
TFTP: 10058752 bytes received in 141 second(s)
File downloaded successfully.
```

```
<IRF> tftp 2.2.2.2 get soft-version2.bin slot2#flash:/soft-version2.bin
```

Specify file **new-config.cfg** as the startup configuration file for all members of the IRF fabric.

```
<IRF> startup saved-configuration new-config.cfg main
Please wait ...
Setting the master board ...
... Done!
Setting the slave board ...
Slot 2:
Set next configuration file successfully
```

Specify file **soft-version2.bin** as the system software image to be used at the next boot of all members of the IRF fabric.

```

<IRF> boot-loader file soft-version2.bin slot all main
  This command will set the boot file of the specified board. Continue? [Y/N]:y
  The specified file will be used as the main boot file at the next reboot on slot
  1!
  The specified file will be used as the main boot file at the next reboot on slot
  2!
# Reboot the switch. The software version is upgraded now.
<IRF> reboot

```

3. Verify the configuration:

Use the **display version** command to check if the upgrade is successful. (Details not shown.)

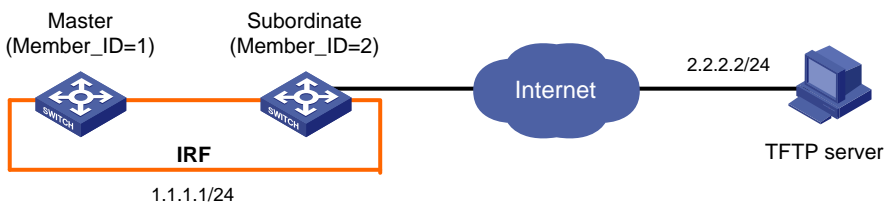
Hotfix configuration example

Network requirements

The IRF fabric in Figure 49 comprises two member switches, the master and subordinate switch. The software running on the member switches has a bug. The patch files **patch_mpu.bin** and **patch_lpu.bin** are saved on the TFTP server. The IRF fabric and TFTP server can reach each other.

From the IRF fabric, use TFTP to download the patch files and then hotfix the software on the fabric.

Figure 49 Network diagram



Note: The orange line represents the IRF link.

Configuration procedure

1. Configure the TFTP server (the configuration varies with server vendors):

- Enable the TFTP server function. (Details not shown.)
- Save the patch files **patch_mpu.bin** and **patch_lpu.bin** to the directory of TFTP server. (Details not shown.)

2. Configure the IRF fabric:

Before upgrading the software, use the **save** command to save the current system configuration. (Details not shown.)

Examine the free space of the Flash on the switch. If the free space is not sufficient for the patch files, delete unused files to release enough space. (Details not shown.)

Load the patch files **patch_mpu.bin** and **patch_lpu.bin** from the TFTP server to the root directory of the master's storage media.

```

<IRF> tftp 2.2.2.2 get patch_mpu.bin
<IRF> tftp 2.2.2.2 get patch_lpu.bin

```

Load the patch files **patch_mpu.bin** and **patch_lpu.bin** from the TFTP server to the root directory of the subordinate switch's storage media.

```

<IRF> tftp 2.2.2.2 get patch_mpu.bin slot2#flash:/patch_mpu.bin
<IRF> tftp 2.2.2.2 get patch_lpu.bin slot2#flash:/patch_lpu.bin

```

Install the patch.

```
<IRF> system-view
```

```
[IRF] patch install flash:
```

```
Patches will be installed. Continue? [Y/N]:y
```

```
Do you want to continue running patches after reboot? [Y/N]:y
```

```
Installing patches.....
```

3. Verify the configuration:

After the installation is complete, use the **display patch information** command to verify whether the patches are installed successfully. (Details not shown.)

Performing ISSU

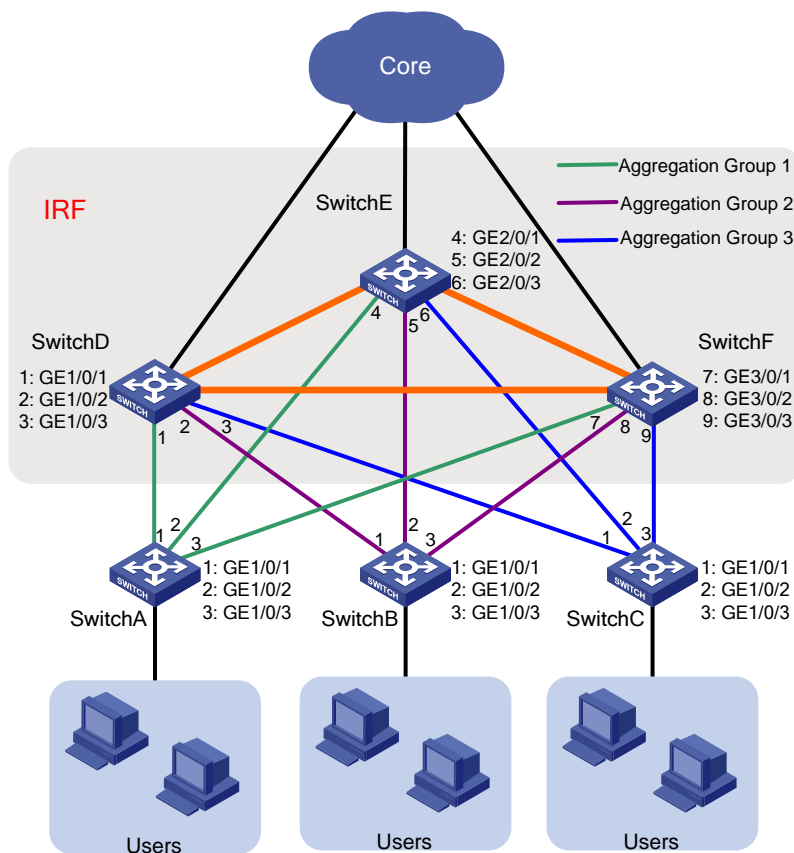
This chapter describes how to use the In-Service Software Upgrade (ISSU) feature to upgrade software. The ISSU feature is available only on the HP 5500-EI switch.

ISSU overview

ISSU enables software upgrade and ensures continuous packet forwarding.

As shown in Figure 50, to ensure high availability for user networks, cross-device link aggregation is configured on the IRF member switches at the distribution layer so that every three physical links with the same color between the IRF member switches and access switches are aggregated as one logical link. In this scenario, you can use ISSU to upgrade system software of each IRF member switch to ensure non-stop forwarding or reduce down time for users connected to Switch A, Switch B, and Switch C.

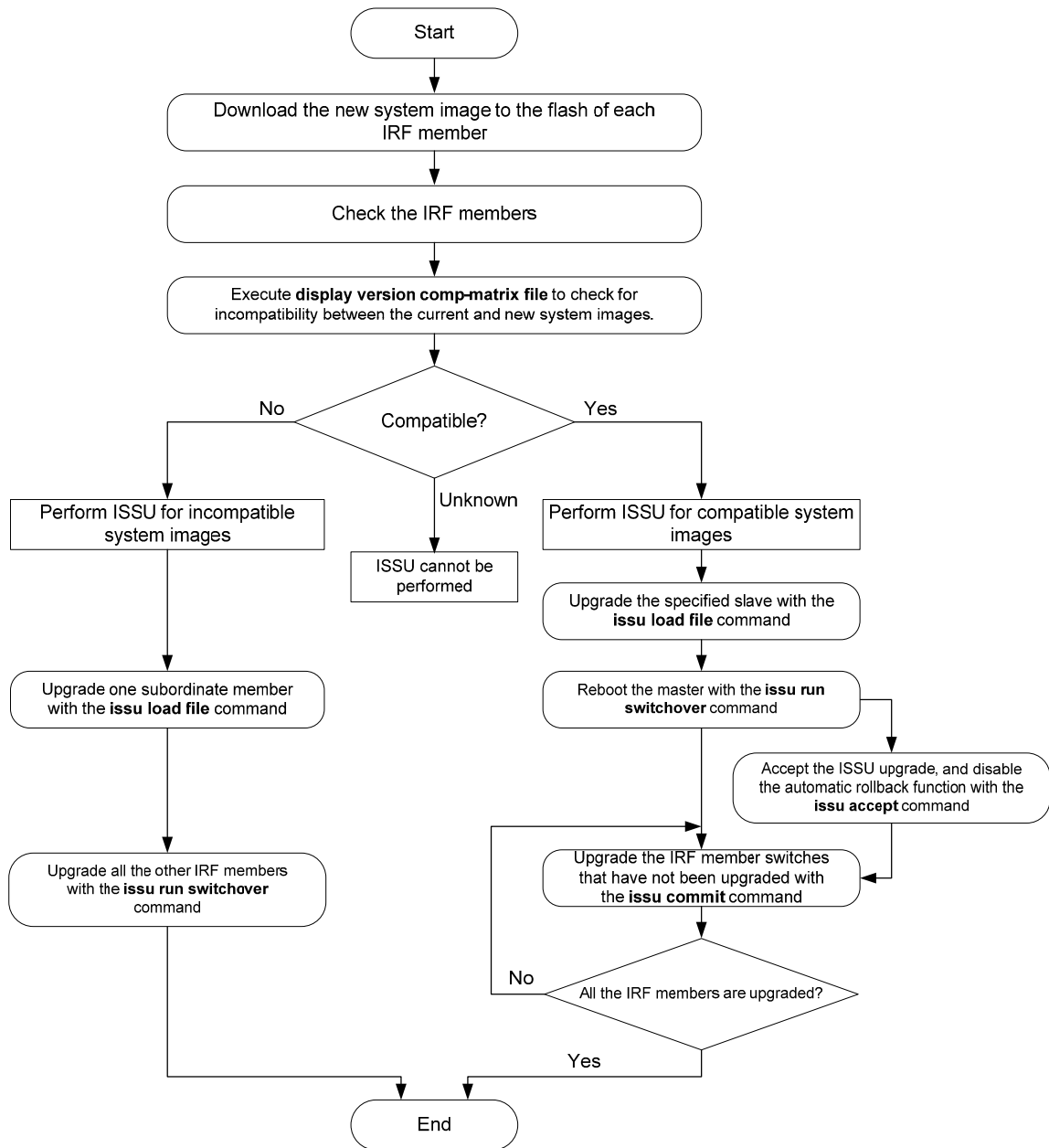
Figure 50 IRF network diagram



ISSU upgrade procedure

ISSU follows a strict procedure, as shown in Figure 51.

Figure 51 ISSU flow chart



IMPORTANT:

- Do not modify the current configuration, plug or unplug cables connected to IRF ports, or delete or modify the system software image during ISSU. Otherwise, the upgrade might fail.
- To upgrade system software of IRF member switches through ISSU, make sure the member switches form a ring topology.

ISSU states

During the ISSU process, you can use the **display issu state** command to display the ISSU state of the IRF fabric, including whether the new system software image is compatible with the current system software image, and the adopted ISSU method.

Table 19 ISSU state description

State	Remarks
Init (Initial state)	No ISSU starts or an ISSU upgrade has completed.
Load	A subordinate switch is being upgraded or has been upgraded. To stop the loading process, perform a manual or automatic roll back to revert system software to its original version.
Switchover	The master is being rebooted to trigger a new master election.
Accept	The ISSU has been accepted. To stop the loading process, you have to perform a manual roll back to revert system software to its original version. The automatic roll back function is not available in this state.
Commit	At least one member switch has not been upgraded to the new version. In this state, neither manual nor automatic rollback can be performed.

System software version rollback

The HP 5500-EI switch series supports version rollback during ISSU. When ISSU fails to proceed on an IRF member switch (for example, the new system software image file is broken), you can use this feature to revert system software to the previous version.

The HP 5500-EI switches support the following version rollback methods.

Automatic rollback

When you reboot the specified subordinate switch with the **issu load** command, the system automatically creates a configurable version rollback timer.

During ISSU for a compatible version, if you do not execute the **issu accept** command on the specified subordinate switch or the **issu commit** command on any other member switch before the rollback timer expires, the system automatically stops the upgrade and rolls back the system software image of the upgraded IRF member switches to the previous version.

During ISSU for an incompatible version, if you do not execute the **issu run switchover** command to upgrade all the IRF member switches that have not been upgraded in one operation before the rollback timer expires, the system automatically rolls back the system software image of all the upgraded IRF member switches to the previous version.

For information about compatible and incompatible ISSU methods, see "[Displaying version compatibility](#)."

Manual rollback

You can use the **issu rollback** command to roll back the system software image of an IRF member switch to the previous version. Whether manual rollback can be performed depends on the ISSU state. For more information, see [Table 19](#).

Performing an ISSU

This section describes how to perform an ISSU.

ISSU upgrade task list

Task	Remarks
Downloading the new system software image to the Flash of all the IRF member switches	Required.
ISSU upgrade prerequisites	Required.
Displaying version compatibility	Required.
Performing an ISSU for a compatible version	Required.
Performing an ISSU for an incompatible version	Use either approach.
<p>If you do not execute the issu run switchover command before the rollback timer expires, the ISSU upgrade ends automatically, the software reverts to the previous version, and the MPUs resume their former roles.</p> <p>The issu run switchover command deletes the rollback timer. After this command is executed, no manual or automatic rollback can be performed for the upgrade.</p>	Optional.
Setting the ISSU version rollback timer	
Performing a manual version rollback	Optional.
Displaying and maintaining ISSU	Optional.

ISSU upgrade prerequisites

Task	Command	Remarks
Save the current configuration.	save	Before performing ISSU, make sure that the current configuration of the IRF fabric has been saved to the configuration file.
Display the running status of each IRF member switch.	display device	Before performing ISSU, make sure that all the member switches of the IRF fabric are in normal state.
Display the information about system software.	display boot-loader	Before performing an ISSU upgrade, make sure that the system software image of all the IRF member switches is identical, which means the version, name, and directory of the system software image are the same.
Display the roles of IRF member switches.	display irf	N/A
Display information about the files in the Flash.	dir	Before performing an ISSU upgrade, make sure that the new and current system software images exist in the Flash of each IRF member switch, and they are saved in the same directory.

Displaying version compatibility

Before performing an ISSU upgrade, check the version compatibility between the new and current system software images, to determine whether ISSU can be performed, and which ISSU method is adopted.

After downloading and saving the new system software image, select an ISSU upgrade method according to one of the following version compatibility check results:

- **Compatible**—The running system software image is compatible with the new system software image. You can use the compatible ISSU method to upgrade system software of the IRF fabric. For more information, see ".".
- **Incompatible**—The current running system software image is incompatible with the new system software image. You can use the incompatible ISSU method to upgrade system software of the IRF fabric. For more information, see "[Performing an ISSU for an incompatible version.](#)"
- **Unknown**—The current and new system software images have big differences, or the current system software image does not support ISSU. You cannot upgrade system software through ISSU.

To display version compatibility:

Step	Command
1. Enters system view.	system-view
2. Check whether the new system software image is compatible with the current system software image.	display version comp-matrix file <i>upgrading-filename</i>

Performing an ISSU for a compatible version

Use the **display version comp-matrix file** *upgrading-filename* command to view the versions of the new and current system software images. If the new system software image is compatible with the current system software image, use this task for ISSU.

ⓘ IMPORTANT:

Before performing compatible ISSU, make sure that the priorities of the current master switch and the specified subordinate switch are higher than other IRF member switches so that the specified subordinate switch can be elected as the new master after the master is rebooted. Otherwise, modify the priorities of the current master switch and the specified subordinate switch with the **irf member** *member-id* **priority** *priority* command.

To perform an ISSU for a compatible version:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Upgrade the specified subordinate switch (the new master after the upgrade).	issu load file <i>upgrading-filename</i> slot <i>slot-number</i>	<i>slot-number</i> is the member ID of the subordinate switch. The specified subordinate switch reboots with the new system software image when this command is executed. Perform the next operation after the subordinate switch is rebooted.

Step	Command	Remarks
3. Reboot the master current manually.	issu run switchover slot <i>slot-number</i>	<p>With this command executed:</p> <ul style="list-style-type: none"> The master reboots with the current system software image, and becomes a subordinate switch after reboot. The subordinate switches of the IRF fabric perform master election. The winner (the subordinate switch specified with the issu load command) becomes the new master. <p>The <i>slot-number</i> provided in this command must be the same as that specified in the issu load command.</p> <p>Perform the next operation after the reboot process is completed.</p>
4. Accept the ISSU.	issu accept slot <i>slot-number</i>	<p>Optional</p> <p>By default, the rollback timer is 45 minutes. If you do not execute the issu accept command on the specified subordinate switch or you do not execute the issu commit command on any other member switch before the rollback timer expires, the system automatically stops the ISSU process and reverts to the previous software version.</p> <p>The <i>slot-number</i> argument provided in this command must be the same as that specified in the issu load command.</p> <p>When this command is executed, the rollback timer becomes invalid and system software cannot be automatically rolled back.</p>
5. Upgrade an IRF member switch that has not been upgraded.	issu commit slot <i>slot-number</i>	<p>This command upgrades one IRF member switch at a time. If the IRF fabric has three or more than three member switches, repeat this command to upgrade them one by one.</p> <p>When all the IRF member switches reboot with the new system software image, the ISSU process completes.</p>

The roles of some IRF member switches change after an ISSU upgrade.

If you do not execute the **issu accept** or **issu commit** command on the specified IRF member switch before the specified rollback timer expires, the system automatically stops the ISSU process and reverts to the previous software version.

After executing the **issu commit** command, you cannot perform version rollback with the **issu rollback** command.

Performing an ISSU for an incompatible version

Use the **display version comp-matrix file** *upgrading-filename* command to view the versions of the new and current system software images. If they are incompatible, use this task for ISSU.

! **IMPORTANT:**

To upgrade system software of the IRF fabric through ISSU when the new and current system software image versions are incompatible, make sure that the multi-active detection (MAD) function has been configured for the IRF fabric. Otherwise, duplicate IRF fabrics will coexist after the new master (the subordinate switch specified with the **issu load** command) is rebooted, causing network faults.

To perform an ISSU for an incompatible version:

Step	Command	Remarks
1. Enters system view.	system-view	N/A
2. Upgrade the specified subordinate switch (the new master after the upgrade).	issu load file <i>upgrading-filename</i> slot <i>slot-number</i> force	<i>slot-number</i> is the member ID of the subordinate switch. After this command is executed, the specified subordinate switch (the new master after the upgrade) reboots with the new system software image, and stays in the Recover state after reboot. Perform the next operation after the reboot process completes.
3. Upgrade all the IRF member switches that have not been upgraded in one operation.	issu run switchover slot <i>slot-number</i>	The <i>slot-number</i> argument provided in this command must be the same as that specified in the issu load command. When this command is executed, all the IRF member switches except the specified subordinate switch (the new master) are upgraded to the new version, and the ISSU process completes after reboot.

If you do not execute the **issu run switchover** command before the rollback timer expires, the ISSU upgrade ends automatically, the software reverts to the previous version, and the MPUs resume their former roles.

The **issu run switchover** command deletes the rollback timer. After this command is executed, no manual or automatic rollback can be performed for the upgrade.

Setting the ISSU version rollback timer

Step	Command	Remarks
1. Enters system view.	system-view	N/A
2. Set the rollback timer.	issu rollback-timer <i>minutes</i>	Optional. By default, the rollback timer is 45 minutes.

During an ISSU upgrade process, if you modify the rollback timer after executing the **issu load** command, the new rollback timer does not take effect for this ISSU process.

Performing a manual version rollback

Step	Command	Remarks
1. Enters system view.	system-view	N/A
2. Perform a manual version rollback.	issu rollback slot <i>slot-number</i>	Optional By default, automatic rollback is performed to revert to the previous version. The <i>slot-number</i> argument provided in this command must be the same as that specified in the issu load command.

Displaying and maintaining ISSU

Task	Command	Remarks
Display information about the rollback timer.	display issu rollback-timer [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the ISSU state.	display issu state	Available in any view
Display version compatibility information.	display version comp-matrix [file <i>upgrading-filename</i>]	Available in any view

ISSU upgrade example

Network status

As shown in [Figure 52](#), access layer switches Switch A, Switch B, and Switch C connect to user networks. Distribution layer switches Switch D, Switch E, and Switch form an IRF fabric. The member ID of the master is 1, and those of the subordinate switches are 2 and 3 respectively.

To ensure high availability, configure cross-device link aggregation using the following guidelines so that every three physical links with the same color between the IRF member switches and access switches are aggregated as one logical link:

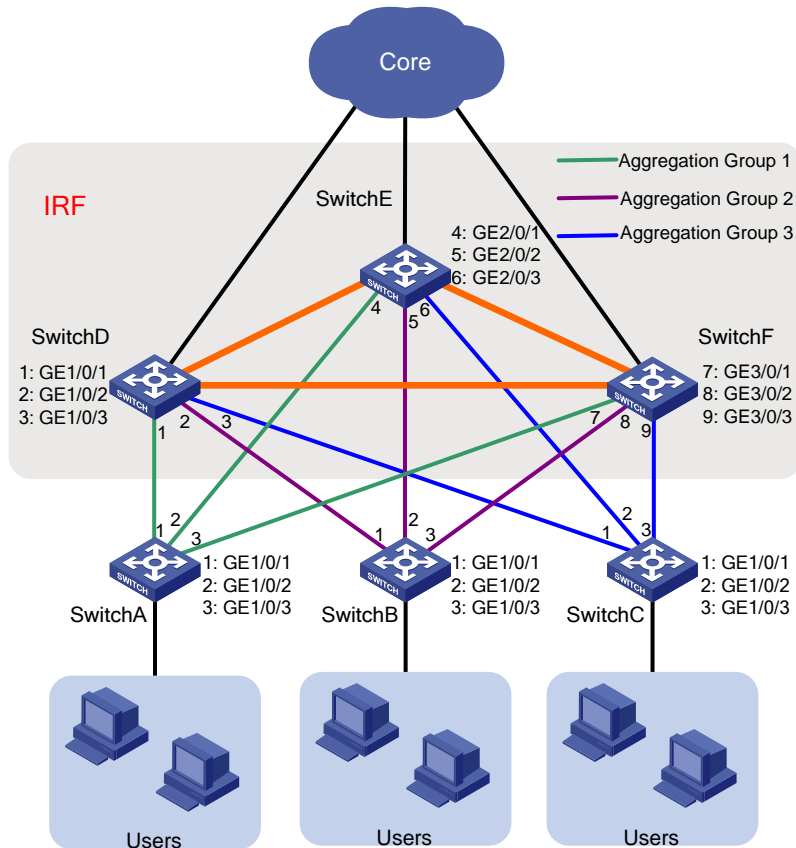
- On the IRF fabric, create three dynamic aggregation groups. Ports in aggregation group 1 connect to Switch A, ports in aggregation group 2 connect to Switch B, and ports in aggregation group 3 connect to Switch C.
- On Switch A, create aggregation group 1 that corresponds to aggregation group 1 on the IRF fabric.
- On Switch B, create aggregation group 2 that corresponds to aggregation group 2 on the IRF fabric.
- On Switch C, create aggregation group 3 that corresponds to aggregation group 3 on the IRF fabric.

Network requirements

The current system software image on each IRF member switch is **soft-version1.bin**. The new system software image **soft-version2.bin** is saved on the TFTP server. The IP address of the IRF fabric is 1.1.1.1/24, and that of the TFTP server is 2.2.2.2/24. The IRF fabric and the TFTP server can reach each other.

Use ISSU to upgrade system software of the IRF member switches to avoid traffic interruption.

Figure 52 Network diagram



Upgrade procedure

Configuring link aggregation

1. Configure the IRF fabric:

Create three dynamic aggregation groups 1, 2, and 3.

```
<IRF> system-view
[IRF] interface bridge-aggregation 1
[IRF-Bridge-Aggregation1] link-aggregation mode dynamic
[IRF-Bridge-Aggregation1] quit
[IRF] interface bridge-aggregation 2
[IRF-Bridge-Aggregation2] link-aggregation mode dynamic
[IRF-Bridge-Aggregation2] quit
[IRF] interface bridge-aggregation 3
[IRF-Bridge-Aggregation3] link-aggregation mode dynamic
[IRF-Bridge-Aggregation3] quit
```

Add ports GigabitEthernet 1/0/1, GigabitEthernet 2/0/1, and GigabitEthernet 3/0/1 that connect to Switch A to aggregation group 1.

```
[IRF] interface GigabitEthernet 1/0/1
[IRF-GigabitEthernet1/0/1] port link-aggregation group 1
[IRF-GigabitEthernet1/0/1] quit
```

```
[IRF] interface GigabitEthernet 2/0/1
[IRF-GigabitEthernet2/0/1] port link-aggregation group 1
[IRF-GigabitEthernet2/0/1] quit
[IRF] interface GigabitEthernet 3/0/1
[IRF-GigabitEthernet3/0/1] port link-aggregation group 1
[IRF-GigabitEthernet3/0/1] quit
```

Add ports GigabitEthernet 1/0/2, GigabitEthernet 2/0/2, and GigabitEthernet 3/0/2 that connect to Switch B to aggregation group 2.

```
[IRF] interface GigabitEthernet 1/0/2
[IRF-GigabitEthernet1/0/2] port link-aggregation group 2
[IRF-GigabitEthernet1/0/2] quit
[IRF] interface GigabitEthernet 2/0/2
[IRF-GigabitEthernet2/0/2] port link-aggregation group 2
[IRF-GigabitEthernet2/0/2] quit
[IRF] interface GigabitEthernet 3/0/2
[IRF-GigabitEthernet3/0/2] port link-aggregation group 2
[IRF-GigabitEthernet3/0/2] quit
```

Add ports GigabitEthernet 1/0/3, GigabitEthernet 2/0/3, and GigabitEthernet 3/0/3 that connect to Switch C to aggregation group 3.

```
[IRF] interface GigabitEthernet 1/0/3
[IRF-GigabitEthernet1/0/3] port link-aggregation group 3
[IRF-GigabitEthernet1/0/3] quit
[IRF] interface GigabitEthernet 2/0/3
[IRF-GigabitEthernet2/0/3] port link-aggregation group 3
[IRF-GigabitEthernet2/0/3] quit
[IRF] interface GigabitEthernet 3/0/3
[IRF-GigabitEthernet3/0/3] port link-aggregation group 3
[IRF-GigabitEthernet3/0/3] quit
```

2. Configure Switch A:

Create dynamic aggregate interface 1.

```
<SwitchA> system-view
[SwitchA] interface bridge-aggregation 1
[SwitchA-Bridge-Aggregation1] link-aggregation mode dynamic
[SwitchA-Bridge-Aggregation1] quit
```

#Add ports GigabitEthernet 1/0/1, GigabitEthernet 1/0/2, and GigabitEthernet 1/0/3 that connect to IRF member switches to aggregation group 1 (corresponding to aggregate interface 1).

```
[SwitchA] interface GigabitEthernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] port link-aggregation group 1
[SwitchA-GigabitEthernet1/0/1] quit
[SwitchA] interface GigabitEthernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] port link-aggregation group 1
[SwitchA-GigabitEthernet1/0/2] quit
[SwitchA] interface GigabitEthernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] port link-aggregation group 1
[SwitchA-GigabitEthernet1/0/3] quit
```

3. Configure Switch B:

Create dynamic aggregate interface 2.

```

<SwitchB> system-view
[SwitchB] interface bridge-aggregation 2
[SwitchB-Bridge-Aggregation2] link-aggregation mode dynamic
[SwitchB-Bridge-Aggregation2] quit

```

#Add ports GigabitEthernet 1/0/1, GigabitEthernet 1/0/2, and GigabitEthernet 1/0/3 that connect to IRF member switches to aggregation group 2 (corresponding to aggregate interface 2).

```

[SwitchB] interface GigabitEthernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] port link-aggregation group 2
[SwitchB-GigabitEthernet1/0/1] quit
[SwitchB] interface GigabitEthernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] port link-aggregation group 2
[SwitchB-GigabitEthernet1/0/2] quit
[SwitchB] interface GigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] port link-aggregation group 2
[SwitchB-GigabitEthernet1/0/3] quit

```

4. Configure Switch C:

Create dynamic aggregate interface 3.

```

<SwitchC> system-view
[SwitchC] interface bridge-aggregation 3
[SwitchC-Bridge-Aggregation3] link-aggregation mode dynamic
[SwitchB-Bridge-Aggregation3] quit

```

Add ports GigabitEthernet 1/0/1, GigabitEthernet 1/0/2, and GigabitEthernet 1/0/3 that connect to IRF member switches to aggregation group 3 (corresponding to aggregate interface 3).

```

[SwitchC] interface GigabitEthernet 1/0/1
[SwitchC-GigabitEthernet1/0/1] port link-aggregation group 3
[SwitchC-GigabitEthernet1/0/1] quit
[SwitchC] interface GigabitEthernet 1/0/2
[SwitchC-GigabitEthernet1/0/2] port link-aggregation group 3
[SwitchC-GigabitEthernet1/0/2] quit
[SwitchC] interface GigabitEthernet 1/0/3
[SwitchC-GigabitEthernet1/0/3] port link-aggregation group 3
[SwitchC-GigabitEthernet1/0/3] quit

```

Configuring the TFTP server

Obtain the new system software image through a legal channel, and save the file to the working directory of the TFTP server so that the TFTP client can access the file. The working directory varies with TFTP server models.

Downloading the new system software image

Download system software image soft-version2.bin to the root directory of the Flash of each IRF member switch.

```

<IRF> tftp 2.2.2.2 get soft-version2.bin
File will be transferred in binary mode
Downloading file from remote TFTP server, please wait.....
TFTP: 10058752 bytes received in 141 second(s)
File downloaded successfully.

```

```
<IRF> copy soft-version2.bin slot2#flash:/
<IRF> copy soft-version2.bin slot3#flash:/
```

Checking all IRF member switches before the ISSU upgrade

1. Check the running status of all IRF member switches. If the running state of a member switch is abnormal, the ISSU upgrade cannot be performed.

```
<IRF> display device
Slot 1
SubSNo PortNum PCBVer FPGAVer CPLDVer BootRomVer AddrLM Type State
0 28 REV.C NULL 002 CC1 IVL MAIN Normal
1 2 REV.A NULL NULL NULL IVL 2*10GE Normal
Slot 2
SubSNo PortNum PCBVer FPGAVer CPLDVer BootRomVer AddrLM Type State
0 28 REV.C NULL 002 CC1 IVL MAIN Normal
2 2 REV.A NULL NULL NULL IVL 2*10GE Normal
Slot 3
SubSNo PortNum PCBVer FPGAVer CPLDVer BootRomVer AddrLM Type State
0 28 REV.C NULL 002 CC1 IVL MAIN Normal
3 2 REV.A NULL NULL NULL IVL 2*10GE Normal
```

The output shows that all IRF member switches are in normal state.

2. Check whether the current system software images on IRF member switches are the same. If not, the ISSU upgrade cannot be performed.

```
<IRF> display boot-loader
Slot 1
The current boot app is: flash:/soft-version1.bin
The main boot app is: flash:/soft-version1.bin
The backup boot app is: flash:/
Slot 2
The current boot app is: flash:/soft-version1.bin
The main boot app is: flash:/soft-version1.bin
The backup boot app is: flash:/
Slot 3
The current boot app is: flash:/soft-version1.bin
The main boot app is: flash:/soft-version1.bin
The backup boot app is: flash:/
```

The output shows that the current system software image file on each IRF member switch is **soft-version1.bin**.

3. View the role of each IRF member switch.

```
<IRF> display irf
Switch Role Priority CPU-Mac Description
*+1 Master 10 0023-8927-ad54 -----
2 Slave 9 0023-8927-afdc -----
3 Slave 1 0023-89d9-3223 -----
-----
```

* indicates the device is the master.

+ indicates the device through which the user logs in.


```
The Bridge MAC of the IRF is: 0023-8927-ad53
Auto upgrade           : yes
Mac persistent        : 6 min
Domain ID             : 0
```

The output shows the following information:

- The member ID and the priority of the master is 1 and 10 respectively
- The member ID and the priority of one subordinate switch is 2 and 9 respectively
- The member ID and the priority of the other subordinate switch is 3 and 1 respectively

During the ISSU upgrade process, you must select subordinate switch 2 as the specified subordinate switch.

Before performing compatible ISSU, make sure the priorities of the master switch and the specified subordinate switch are higher than other IRF member switches so that the specified subordinate switch can be elected as the new master after the master is rebooted. Otherwise, modify the priorities of the master switch and the specified subordinate switch with the **irf member member-id priority priority** command.

4. Check whether the new system software image has been saved in the Flash of each IRF member switch. If not, the ISSU upgrade cannot be performed.

Verify whether the new system software image **soft-version2.bin** has been saved to the Flash of the master.

```
<IRF> dir
```

```
Directory of flash:/
```

0	-rw-	6085	May 29 2010 11:38:45	config.cfg
1	-rw-	10518	Apr 26 2011 12:45:05	logfile.log
2	-rw-	12397691	Apr 26 2011 14:24:11	soft-version1.bin
3	-rw-	13308645	Apr 26 2011 14:13:46	soft-version2.bin
4	drw-	-	Apr 26 2011 12:00:33	seclog
5	-rw-	287	Apr 26 2011 12:19:52	system.xml

```
31496 KB total (5981 KB free)
```

The output shows that the new system software image has been saved to the Flash of the master.

Verify whether the new system software image **soft-version2.bin** has been saved to the Flash of subordinate switch 2.

```
<IRF> dir slot2#flash:/
```

```
Directory of slot2#flash:/
```

0	-rw-	6085	May 29 2010 11:38:45	config.cfg
1	-rw-	10518	Apr 26 2011 12:45:05	logfile.log
2	-rw-	12397691	Apr 26 2011 14:24:11	soft-version1.bin
3	-rw-	13308645	Apr 26 2011 14:13:46	soft-version2.bin
4	drw-	-	Apr 26 2011 12:00:33	seclog
5	-rw-	287	Apr 26 2011 12:19:52	system.xml

```
31496 KB total (5981 KB free)
```

The output shows that the new system software image has been saved to the Flash of subordinate switch 2.

Verify whether the new system software image **soft-version2.bin** has been saved to the Flash of subordinate switch 3.

```
<IRF> dir slot3#flash:/
Directory of slot3#flash:/

 0  -rw-     6085  May 29 2010 11:38:45  config.cfg
 1  -rw-    10518  Apr 26 2011 12:45:05  logfile.log
 2  -rw- 12397691  Apr 26 2011 14:24:11  soft-version1.bin
 3  -rw- 13308645  Apr 26 2011 14:13:46  soft-version2.bin
 4  drw-      -   Apr 26 2011 12:00:33  seclog
 5  -rw-     287  Apr 26 2011 12:19:52  system.xml
```

31496 KB total (5981 KB free)

The output shows that the new system software image has been saved to the Flash of subordinate switch 3.

5. Save the current configuration.

```
<IRF> save
The current configuration will be written to the device. Are you sure? [Y/N]:y
Please input the file name(*.cfg)[flash:/config.cfg]
(To leave the existing filename unchanged, press the enter key):
flash:/config.cfg exists, overwrite? [Y/N]:y
Validating file. Please wait....
Saved the current configuration to mainboard device successfully.
Slot 2:
Save next configuration file successfully.
Slot 3:
Save next configuration file successfully.
Configuration is saved to device successfully.
```

The output shows that the current configuration has been saved to the configuration file of each IRF member switch.

Viewing the version compatibility

Check whether the current and new system software images are compatible.

```
<IRF> display version comp-matrix file soft-version2.bin
```

- If they are compatible, the output is like the following:

```
Number of Matrices in Table = 1
Matrix for HP 5500-EI
```

```
Running Version: version1
Version Compatibility List:
version2 (Compatible)
```

The output shows that the new and current versions are fully compatible. You can use the compatible ISSU upgrade method. For more information, see "[Performing compatible ISSU upgrade.](#)"

- If the two versions are incompatible, the output is like the following:

```
Number of Matrices in Table = 1
```

Matrix for HP 5500-EI

```
Running Version: version1
Version Compatibility List:
version2 (Incompatible)
```

The output shows that the two versions are incompatible. You must use the incompatible ISSU method. For more information, see "[Performing incompatible ISSU upgrade.](#)"

Performing compatible ISSU upgrade

Upgrade the specified subordinate switch (the new master after the upgrade), which is subordinate switch 2 in this example.

```
<IRF> system-view
```

```
[IRF] issu load file soft-version2.bin slot 2
```

This command will begin ISSU, and the specified board will reboot and be upgraded. Please save the current running configuration first; otherwise, the configuration may be lost. Continue? [Y/N]:y

After the reboot of subordinate switch 2, check whether the system software image of slave switch 2 is **soft-version2.bin**.

```
[IRF] display boot-loader
```

```
Slot 1
```

```
The current boot app is: flash:/soft-version1.bin
```

```
The main boot app is: flash:/soft-version1.bin
```

```
The backup boot app is: flash:/
```

```
Slot 2
```

```
The current boot app is: flash:/soft-version2.bin
```

```
The main boot app is: flash:/soft-version2.bin
```

```
The backup boot app is: flash:/
```

```
Slot 3
```

```
The current boot app is: flash:/soft-version1.bin
```

```
The main boot app is: flash:/soft-version1.bin
```

```
The backup boot app is: flash:/
```

The output shows that the system software image of subordinate switch 2 is soft-version2.bin.

Reboot the master manually.

```
[IRF] issu run switchover slot 2
```

Master will reboot, switch the specified board to master and update the line card. Continue? [Y/N]:y

In this example, the member ID of the master is 1. After reboot, the master becomes a subordinate switch in the IRF fabric. Then the subordinate switches perform a role election, and slave 2 becomes the new master because the priority of subordinate switch 2 is higher than that of subordinate switch 3.

Accept the ISSU upgrade and delete the rollback timer.

```
[IRF] issu accept slot 2
```

Upgrade switch 1 and switch 3.

```
[IRF] issu commit slot 1
```

The specified board will reboot and be upgraded. Continue? [Y/N]:y

```
[IRF] issu commit slot 3
```

The specified board will reboot and be upgraded. Continue? [Y/N]:y

Then the ISSU upgrade process completes and the system software images of all IRF member switches have been upgraded to the new version.

Verify whether the current system software images on the IRF member switches are **soft-version2.bin**.

```
[IRF] display boot-loader
```

```
Slot 1
```

```
The current boot app is: flash:/soft-version2.bin
```

```
The main boot app is: flash:/soft-version2.bin
```

```
The backup boot app is: flash:/
```

```
Slot 2
```

```
The current boot app is: flash:/soft-version2.bin
```

```
The main boot app is: flash:/soft-version2.bin
```

```
The backup boot app is: flash:/
```

```
Slot 3
```

```
The current boot app is: flash:/soft-version2.bin
```

```
The main boot app is: flash:/soft-version2.bin
```

```
The backup boot app is: flash:/
```

Performing incompatible ISSU upgrade



IMPORTANT:

To upgrade system software of the IRF fabric through ISSU when the new and current system software image versions are incompatible, make sure that the multi-active detection (MAD) function has been configured for the IRF fabric. Otherwise, duplicate IRF fabrics will coexist after the new master (the subordinate switch specified with the **issu load** command) is rebooted, causing network faults.

Upgrade the specified subordinate switch (the new master after the upgrade), which is subordinate switch 2 in this example.

```
<IRF> system-view
```

```
[IRF] issu load file soft-version2.bin slot 2
```

```
This command will begin ISSU, and the specified board will reboot and be upgraded. Please save the current running configuration first; otherwise, the configuration may be lost.Continue? [Y/N]: y
```

After the reboot of subordinate switch 2, reboot and upgrade all IRF member switches that have not been upgraded.

```
[IRF] issu run switchover slot 2
```

```
Master will reboot, switch the specified board to master and update the line card. Continue? [Y/N]:y
```

Then, the ISSU upgrade process completes and the system software images of all IRF member switches have been upgraded to the new version.

Verify whether the current system software images on the IRF member switches are **soft-version2.bin**.

```
[IRF] display boot-loader
```

```
Slot 1
```

```
The current boot app is: flash:/soft-version2.bin
```

```
The main boot app is: flash:/soft-version2.bin
```

```
The backup boot app is: flash:/
```

```
Slot 2
```

```
The current boot app is: flash:/soft-version2.bin
```

```
The main boot app is:    flash:/soft-version2.bin
The backup boot app is: flash:/
Slot 3
The current boot app is: flash:/soft-version2.bin
The main boot app is:    flash:/soft-version2.bin
The backup boot app is:  flash:/
```

Managing the device

Device management includes monitoring the operating status of devices and configuring their running parameters.

The configuration tasks in this document are order independent. You can perform these tasks in any order.

Configuring the device name

A device name identifies a device in a network and works as the user view prompt at the CLI. For example, if the device name is **Sysname**, the user view prompt is <Sysname>.

To configure the device name:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the device name.	sysname <i>sysname</i>	Optional. The default device name is HP .

Changing the system time

You must synchronize your device with a trusted time source by using NTP or changing the system time before you run it on the network. Network management depends on an accurate system time setting, because the timestamps of system messages and logs use the system time.

In a small-sized network, you can manually set the system time of each device.

Configuration guidelines

You can change the system time by configuring the relative time, time zone, and daylight saving time. The configuration result depends on their configuration order (see [Table 20](#)). In the first column of this table, 1 represents the **clock datetime** command, 2 represents the **clock timezone** command, and 3 represents the **clock summer-time** command. To verify the system time setting, use the **display clock** command. This table assumes that the original system time is 2005/1/1 1:00:00.

Table 20 System time configuration results

Command	Effective system time	Configuration example	System time
1	<i>date-time</i>	<code>clock datetime 1:00 2007/1/1</code>	01:00:00 UTC Mon 01/01/2007.
2	Original system time ± <i>zone-offset</i>	<code>clock timezone zone-time add 1</code>	02:00:00 zone-time Sat 01/01/2005.
1, 2	<i>date-time</i> ± <i>zone-offset</i>	<code>clock datetime 2:00 2007/2/2</code> <code>clock timezone zone-time add 1</code>	03:00:00 zone-time Fri 02/02/2007.

Command	Effective system time	Configuration example	System time
2, 1	<i>date-time</i>	clock timezone zone-time add 1 clock datetime 3:00 2007/3/3	03:00:00 zone-time Sat 03/03/2007.
	The original system time outside the daylight saving time range: The system time does not change until it falls into the daylight saving time range.	clock summer-time ss one-off 1:00 2006/1/1 1:00 2006/8/8 2	01:00:00 UTC Sat 01/01/2005.
3	The original system time in the daylight saving time range: The system time increases by <i>summer-offset</i> .	clock summer-time ss one-off 00:30 2005/1/1 1:00 2005/8/8 2	03:00:00 ss Sat 01/01/2005. NOTE: If the original system time plus <i>summer-offset</i> is beyond the daylight saving time range, the original system time does not change. After you disable the daylight saving setting, the system time automatically decreases by <i>summer-offset</i> .
	<i>date-time</i> outside the daylight saving time range: <i>date-time</i>	clock datetime 1:00 2007/1/1 clock summer-time ss one-off 1:00 2006/1/1 1:00 2006/8/8 2	01:00:00 UTC Mon 01/01/2007.
1, 3	<i>date-time</i> in the daylight saving time range: <i>date-time</i> + <i>summer-offset</i>	clock datetime 8:00 2007/1/1 clock summer-time ss one-off 1:00 2007/1/1 1:00 2007/8/8 2	10:00:00 ss Mon 01/01/2007. NOTE: If the <i>date-time</i> plus <i>summer-offset</i> is outside the daylight saving time range, the system time equals <i>date-time</i> . After you disable the daylight saving setting, the system time automatically decreases by <i>summer-offset</i> .
3, 1 (<i>date-time</i> outside the daylight saving time range)	<i>date-time</i>	clock summer-time ss one-off 1:00 2007/1/1 1:00 2007/8/8 2 clock datetime 1:00 2008/1/1	01:00:00 UTC Tue 01/01/2008.

Command	Effective system time	Configuration example	System time
3, 1 (<i>date-time</i> in the daylight saving time range)	<i>date-time</i> – <i>summer-offset</i> outside the daylight saving time range:	clock summer-time ss one-off 1:00 2007/1/1 1:00 2007/8/8 2	23:30:00 UTC Sun 12/31/2006.
	<i>date-time</i> – <i>summer-offset</i>	clock datetime 1:30 2007/1/1	
	<i>date-time</i> – <i>summer-offset</i> in the daylight saving time range:	clock summer-time ss one-off 1:00 2007/1/1 1:00 2007/8/8 2	03:00:00 ss Mon 01/01/2007.
	<i>date-time</i>	clock datetime 3:00 2007/1/1	
2, 3 or 3, 2	Original system clock ± <i>zone-offset</i> outside the daylight saving time range:	clock timezone zone-time add 1 clock summer-time ss one-off 1:00 2007/1/1 1:00 2007/8/8 2	02:00:00 zone-time Sat 01/01/2005.
	Original system clock ± <i>zone-offset</i>		
	Original system clock ± <i>zone-offset</i> outside the daylight saving time range:	clock timezone zone-time add 1 clock summer-time ss one-off 1:00 2005/1/1 1:00 2005/8/8 2	System clock configured: 04:00:00 ss Sat 01/01/2005.
	Original system clock ± <i>zone-offset</i> + <i>summer-offset</i>		
1, 2, 3 or 1, 3, 2	<i>date-time</i> ± <i>zone-offset</i> outside the daylight saving time range:	clock datetime 1:00 2007/1/1 clock timezone zone-time add 1 clock summer-time ss one-off 1:00 2008/1/1 1:00 2008/8/8 2	02:00:00 zone-time Mon 01/01/2007.
	<i>date-time</i> ± <i>zone-offset</i>	clock datetime 1:00 2007/1/1	
	<i>date-time</i> ± <i>zone-offset</i> outside the daylight saving time range:	clock timezone zone-time add 1 clock summer-time ss one-off 1:00 2007/1/1 1:00 2007/8/8 2	04:00:00 ss Mon 01/01/2007.
	<i>date-time</i> ± <i>zone-offset</i> + <i>summer-offset</i>		
2, 3, 1 or 3, 2, 1	<i>date-time</i> outside the daylight saving time range:	clock timezone zone-time add 1 clock summer-time ss one-off 1:00 2008/1/1 1:00 2008/8/8 2	01:00:00 zone-time Mon 01/01/2007.
	<i>date-time</i>	clock datetime 1:00 2007/1/1	

Command	Effective system time	Configuration example	System time
	<i>date-time</i> in the daylight saving time range, but <i>date-time – summer-offset</i> outside the summer-time range:	<pre>clock timezone zone-time add 1 clock summer-time ss one-off 1:00 2008/1/1 1:00 2008/8/8 2</pre>	23:30:00 zone-time Mon 12/31/2007.
	<i>date-time – summer-offset</i>	<pre>clock datetime 1:30 2008/1/1</pre>	
	Both <i>date-time</i> and <i>date-time – summer-offset</i> in the daylight saving time range:	<pre>clock timezone zone-time add 1 clock summer-time ss one-off 1:00 2008/1/1 1:00 2008/8/8 2</pre>	03:00:00 ss Tue 01/01/2008.
	<i>date-time</i>	<pre>clock datetime 3:00 2008/1/1</pre>	

Configuration procedure

To change the system time:

Step	Command	Remarks
1. Set the system time and date.	clock datetime <i>time date</i>	Optional. Available in user view.
2. Enter system view.	system-view	N/A
3. Set the time zone.	clock timezone <i>zone-name</i> { add minus } <i>zone-offset</i>	Optional. Universal time coordinated (UTC) time zone by default.
4. Set a daylight saving time scheme.	<ul style="list-style-type: none"> Set a non-recurring scheme: clock summer-time <i>zone-name one-off start-time start-date end-time end-date add-time</i> Set a recurring scheme: clock summer-time <i>zone-name repeating start-time start-date end-time end-date add-time</i> 	Optional. Use either command. By default, daylight saving time is disabled, and the UTC time zone applies.

Enabling displaying the copyright statement

The device by default displays the copyright statement when a Telnet or SSH user logs in, or when a console user quits user view. You can disable or enable the function as needed. The following is a sample copyright statement:

```
*****
* Copyright (c) 2010-2011 Hewlett-Packard Development Company, L.P.          *
* Without the owner's prior written consent,                                *
* no decompiling or reverse-engineering shall be allowed.                  *
*****
```

To enable displaying the copyright statement:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable displaying the copyright statement.	copyright-info enable	Optional Enabled by default.

Changing the brand name

An IRF fabric using HP, H3C, and 3Com member switches together might cause network management problems. For example, if an IRF fabric uses an HP switch as the master member switch and uses an H3C switch as a subordinate member switch, the network management software identifies the switch as an HP IRF fabric. Then, if the IRF fabric is rebooted and the H3C switch is elected as the master, the network management software identifies the IRF fabric as a new IRF fabric. To avoid this problem, perform this task to change the brand names of all member switches to the same.

Table 21 lists the model matrix for HP, H3C, and 3Com switches. You can change brand names among them, except the models marked N/A. Before you change the brand name, you must get clear of the matrix for proper operation of the switches.

Table 21 Brand matrix

HP switch model	H3C switch model	3COM switch model
HP 5500-24G EI Switch with 2 Interface Slots	S5500-28C-EI	Switch 4800G 24-Port
HP 5500-48G EI Switch with 2 Interface Slots	S5500-52C-EI	Switch 4800G 48-Port
HP 5500-24G-SFP EI Switch with 2 Interface Slots	S5500-28F-EI	Switch 4800G 24-Port SFP
HP 5500-24G-PoE+ EI Switch with 2 Interface Slots	S5500-28C-PWR-EI	Switch 4800G PWR 24-Port
HP 5500-48G-PoE+ EI Switch with 2 Interface Slots	S5500-52C-PWR-EI	Switch 4800G PWR 48-Port
HP 5500-24G SI Switch with 2 Interface Slots	S5500-28C-SI	N/A
HP 5500-48G SI Switch with 2 Interface Slots	S5500-52C-SI	N/A
HP 5500-24G-PoE+ SI Switch with 2 Interface Slots	S5500-28C-PWR-SI	N/A
HP 5500-48G-PoE+ SI Switch with 2 Interface Slots	S5500-52C-PWR-SI	N/A

Configuration preparation

Before you change the brand name for an HP, H3C, or 3Com switch, prepare the proper Boot ROM and system software image file according to the model matrix as listed in Table 21. The following describes

the procedure for changing the brand name of an H3C or 3Com switch to HP. The procedure is the same for changing the brand names among HP, H3C, and 3Com switches.

1. Load the proper HP Boot ROM to the flash memory of the H3C or 3Com switch and use the HP Boot ROM to upgrade the Boot ROM of the switch.
2. Load the proper HP system software image file to the flash memory of the H3C or 3Com switch, specify the file as the main system software image file, and reboot the switch.
3. Execute the **brand** command and reboot the switch.

NOTE:

For HP 5500 EI and HP 5500SI use the **bootrom update** command to upgrade the Boot ROM.

Configuration guidelines

- After you change the brand name of a 3Com switch to HP or H3C, the default baudrate of the console port changes from 19200 to 9600.
- The port numbering rule for 3Com switches is different from that for HP and H3C switches. After you change the brand name for a 3Com switch, the port numbers become inconsistent with the silkscreen marks. Configure the port (if necessary) according to the numbering rules for the models in the model matrix in Table 21.
- The default settings for some features on 3Com switches are different from those on HP and H3C switches. After you change the brand name for a switch, the default settings for those features become the default settings of the target brand.

Configuration procedure

You can use the **display brand** command to display the brand names of the member switches. If any consistent brand names exist in the IRF fabric, change them to the same.

To change brand name for a member switch:

Step	Command
1. Change the brand name for a member switch.	brand { hp h3c 3com } [slot slot-number]
2. Reboot the member switch.	reboot slot slot-number

After you change the brand name for a member switch, the switch can use the later software versions for the new brand.

NOTE:

The default settings vary with different brands. Changing the brand name might affect the running configuration. After you change the brand name of a member switch, verify the configuration and re-configure the switch if necessary.

Configuring banners

Banners are messages that the system displays during user login.

The system supports the following banners:

- **Legal banner**—Appears after the copyright or license statement. To continue login, the user must enter **Y** or press **Enter**. To quit the process, the user must enter **N**. **Y** and **N** are case-insensitive.
- **Message of the Day (MOTD) banner**—Appears after the legal banner and before the login banner.
- **Login banner**—Appears only when password or scheme authentication has been configured.
- **Incoming banner**—Appears for Modem users.
- **Shell banner**—Appears for non-Modem users.

Banner message input modes

You can configure a banner in one of the following ways:

- Single-line input

Input the entire banner in the same line as the command. The start and end delimiters for the banner must be the same but can be any visible character. The input text, including the command keywords and the delimiters cannot exceed 510 characters. In this mode, do not press **Enter** before you input the end delimiter. For example, you can configure the shell banner “Have a nice day.” as follows:

```
<System> system-view
[System] header shell %Have a nice day.%
```

- Multiple-line input

Input message text in multiple lines. In this approach, the message text can be up to 2000 characters. Use one of the following methods to implement multi-line input mode:

- **Method 1**—Press **Enter** after the last command keyword. At the system prompt, enter the banner message and end with the delimiter character **%**. For example, you can configure the banner “Have a nice day. Please input the password.” as follows:

```
<System> system-view
[System] header shell
Please input banner content, and quit with the character '%'.--System prompt
Have a nice day.
Please input the password.%
```

- **Method 2**—After you type the last command keyword, type any character as the start delimiter for the banner message and press **Enter**. At the system prompt, type the banner message and end the last line with a delimiter that is the same as the start delimiter. For example, you can configure the banner “Have a nice day. Please input the password.” as follows:

```
<System> system-view
[System] header shell A
Please input banner content, and quit with the character 'A'.--System prompt
Have a nice day.
Please input the password.A
```

- **Method 3**—After you type the last keyword, type the start delimiter and part of the banner message and press **Enter**. At the system prompt, enter the rest of the banner and end the last line with a delimiter that is the same as the start delimiter. In this approach, you can use any character as the start and end delimiters but must make sure that it is not the same as the end character of the message text in the first line. For example, you can configure the banner “Have a nice day. Please input the password.” as follows:

```
<System> system-view
[System] header shell AHave a nice day.
```

```
Please input banner content, and quit with the character 'A'.--System prompt
Please input the password.A
```

Configuration procedure

To configure a banner:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the incoming banner.	header incoming <i>text</i>	Optional
3. Configure the login banner.	header login <i>text</i>	Optional
4. Configure the legal banner.	header legal <i>text</i>	Optional
5. Configure the shell banner.	header shell <i>text</i>	Optional
6. Configure the MOTD banner.	header motd <i>text</i>	Optional

Configuring the exception handling method

You can configure the device to handle system exceptions in one of the following methods:

- **reboot**—The device automatically reboots to recover from the error condition.
- **maintain**—The device stays in the error condition so you can collect complete data, including error messages, for diagnosis. In this approach, you must manually reboot the device.

To configure the exception handling method:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the exception handling method.	system-failure { maintain reboot }	Optional. By default, the system reboots when an exception occurs.

NOTE:

In an IRF fabric, the exception handling method applies to all member switches, but the member switches handle system exceptions independently without affecting one another.

Rebooting the device

△ CAUTION:

- A reboot can interrupt network services.
 - To avoid data loss, use the **save** command to save the current configuration before a reboot.
 - Use the **display startup** and **display boot-loader** commands to verify that you have correctly set the startup configuration file and the main system software image file. If the main system software image file has been corrupted or does not exist, the device cannot reboot. You must re-specify a main system software image file, or power off the device and then power it on so the system can reboot with the backup system software image file.
-

You can reboot the device in one of the following ways to recover from an error condition:

- Reboot the device immediately at the CLI.
- At the CLI, schedule a reboot to occur at a specific time and date or after a delay.
- Power off and then power on the device. This method might cause data loss and hardware damage, and is the least preferred method.
- Reboot at the CLI enables easy remote device maintenance.

Rebooting devices immediately at the CLI

To reboot a device, perform the following task in user view:

Task	Command	Remarks
Reboot a switch or all IRF member switches immediately.	reboot [<i>slot slot-number</i>]	If you do not specify any IRF member ID for the <i>slot-number</i> argument, all IRF member switches reboot.

Scheduling a device reboot

The switch supports only one device reboot schedule. If you configure the **schedule reboot delay** command multiple times, the last configuration takes effect.

The **schedule reboot at** command and the **schedule reboot delay** command overwrite each other, and whichever is configured last takes effect.

For data security, if you are performing file operations at the reboot time, the system does not reboot.

To schedule a device reboot, perform the following task in user view:

Task	Command	Remarks
Schedule a reboot.	<ul style="list-style-type: none">• Schedule a reboot to occur at a specific time and date: schedule reboot at <i>hh:mm</i> [<i>date</i>]• Schedule a reboot to occur after a delay: schedule reboot delay { <i>hh:mm</i> <i>mm</i> }	<p>Use either command.</p> <p>The scheduled reboot function is disabled by default.</p> <p>Changing any clock setting can cancel the reboot schedule.</p>

Scheduling jobs

You can schedule a job to automatically run a command or a set of commands without administrative interference. The commands in a job are polled every minute. When the scheduled time for a command is reached, the job automatically executes the command. If a confirmation is required while the command is running, the system automatically inputs **Y** or **Yes**. If characters are required, the system automatically inputs a default character string, or inputs an empty character string when there is no default character string.

Job configuration approaches

You can configure jobs in a non-modular or modular approach. Use the non-modular approach for a one-time command execution and use non-modular approach for complex maintenance work.

Table 22 A comparison of non-modular and modular approaches

Comparison item	Scheduling a job in the non-modular approach	Scheduling a job in the modular approach
Configuration method	Configure all elements in one command	Separate job, view, and time settings.
Can multiple jobs be configured?	No	Yes
Can a job have multiple commands?	No If you use the schedule job command repeatedly, only the last configuration takes effect.	Yes You can use the time command in job view to configure commands to be executed at different time points.
Supported views	User view and system view. In the schedule job command, shell represents user view, and system represents system view.	All views. In the time command, monitor represents user view.
Supported commands	Commands in user view and system view	Commands in all views.
Can a job be repeatedly executed?	No	Yes
Can a job be saved to the configuration file?	No	Yes

Configuration guidelines

- To have a job successfully run a command, check that the specified view and command are valid. The system does not verify their validity.
- The configuration interface, view, and user status that you have before job execution restores even if the job has run a command that changes the user interface (for example, **telnet**, **ftp**, and **ssh2**), the view (for example, **system-view** and **quit**), or the user status (for example, **super**).
- The jobs run in the background without displaying any messages except log, trap and debugging messages.
- In the modular approach:
 - Every job can have only one view and up to 10 commands. If you specify multiple views, the one specified the last takes effect.

- Input a view name in its complete form. Most commonly used view names include **monitor** for user view, **system** for system view, **GigabitEthernetx/x/x**, and **Ten-GigabitEthernetx/x/x** for Ethernet interface view, and **Vlan-interfacex** for VLAN interface view.
- The time ID (*time-id*) must be unique in a job. If two time and command bindings have the same time ID, the one configured last takes effect.

Scheduling a job in the non-modular approach

Perform one of the following commands in user view to schedule a job:

Step	Command	Remarks
Schedule a job.	<ul style="list-style-type: none"> • Schedule a job to run a command at a specific time: schedule job at time [date] view view command • Schedule a job to run a command after a delay: schedule job delay time view view command 	Use either command. NOTE: <ul style="list-style-type: none"> • If you execute the schedule job command repeatedly, the last configuration takes effect. • Changing any clock setting can cancel the job set by using the schedule job command.

Scheduling a job in the modular approach

To configure a scheduled job:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a job and enter job view.	job job-name	N/A
3. Specify the view in which the commands in the job run.	view view-name	You can specify only one view for a job. The job executes all commands in the specified view.
4. Add commands to the job.	<ul style="list-style-type: none"> • Configure a command to run at a specific time and date: time time-id at time date command command • Configure a command to run at a specific time: time time-id { one-off repeating } at time [month-date month-day week-day week-daylist] command command • Configure a command to run after a delay: time time-id { one-off repeating } delay time command command 	Use any of the commands. NOTE: Changing a clock setting does not affect the schedule set by using the time at or time delay command.

Disabling Boot ROM access

By default, anyone can press **Ctrl+B** during startup to enter the Boot menu and configure the Boot ROM. To protect the system, you can disable Boot ROM access so the users can access only the CLI.

You can also set a Boot ROM password the first time you access the Boot menu to protect the Boot ROM.

To view Boot ROM accessibility status, use the **display startup** command. For more information about the **display startup** command, see *Fundamentals Command Reference*.

Follow the step below to disable Boot ROM access:

Task	Command	Remarks
Disable Boot ROM access.	undo startup bootrom-access enable	By default, Boot ROM access is enabled. Available in user view.

Configuring the port status detection timer

Some protocols might shut down ports under specific circumstances. For example, MSTP shuts down a BPDU guard enabled port when the port receives a BPDU. Then, the device starts the detection timer. If the port is still down when the detection timer expires, the port quits the shutdown status and resumes its actual physical status.

To configure the port status detection timer:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the port status detection timer.	shutdown-interval time	Optional. The detection timer is 30 seconds by default.

Configuring temperature thresholds for a device

You can set the temperature thresholds to monitor the temperature of a device.

The temperature thresholds include lower threshold and warning threshold.

When the device temperature drops below the lower threshold or reaches the warning threshold, the device logs the event and outputs a log message and a trap.

When the device temperature reaches the alarming threshold, the device logs the event and outputs a log message and a trap repeatedly.

When the device temperature reaches the shut-down threshold, the device logs the event, outputs a log message and a trap, and automatically shuts down.

To configure temperature thresholds for an IRF member device:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure temperature thresholds for an IRF member device.	temperature-limit slot <i>slot-number</i> hotspot <i>sensor-number</i> lowerlimit <i>warninglimit</i>	Optional. By default, the lower threshold is -5°C (23°F), and the warning threshold is 55°C (131°F). The warning threshold must be higher than the lower threshold.

Clearing unused 16-bit interface indexes

The device must maintain persistent 16-bit interface indexes and keep one interface index match one interface name for network management. After deleting a logical interface, the device retains its 16-bit interface index so the same index can be assigned to the interface at interface re-creation.

To avoid index depletion causing interface creation failures, you can clear all 16-bit indexes that have been assigned but not in use. The operation does not affect the interface indexes of the interfaces that have been created but the indexes assigned to re-created interfaces might change.

! IMPORTANT:

A confirmation is required when you execute this command. The command will not run if you fail to make a confirmation within 30 seconds or enter **N** to cancel the operation.

To clear unused 16-bit interface indexes, perform the following task in user view:

Task	Command	Remarks
Clear unused 16-bit interface indexes.	reset unused porttag	In an IRF fabric, the command applies to all member switches.

Verifying and diagnosing transceiver modules

You can verify the genuineness of a transceiver module in the following ways:

- Display the key parameters of a transceiver module, including its transceiver type, connector type, central wavelength of the transmit laser, transfer distance and vendor name.
- Display its electronic label. The electronic label is a profile of the transceiver module and contains the permanent configuration including the serial number, manufacturing date, and vendor name. The data is written to the storage component during debugging or testing.

To verify transceiver modules, perform the following tasks in any view:

Task	Command
Display key parameters of transceiver modules.	display transceiver interface [<i>interface-type</i> <i>interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]

Task	Command
Display electronic label data for transceiver modules.	display transceiver manuinfo interface [<i>interface-type</i> <i>interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]

Diagnosing transceiver modules

The device provides the alarm function and digital diagnosis function for transceiver modules. When a transceiver module fails or inappropriately operates, you can check for alarms present on the transceiver module to identify the fault source or examine the key parameters monitored by the digital diagnosis function, including the temperature, voltage, laser bias current, TX power, and RX power.

To diagnose transceiver modules, perform the following tasks in any view:

Task	Command
Display alarms present on transceiver modules.	display transceiver alarm interface [<i>interface-type</i> <i>interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]
Display the present measured values of the digital diagnosis parameters for pluggable transceivers.	display transceiver diagnosis interface [<i>interface-type</i> <i>interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]

Displaying and maintaining device management

For diagnosis or troubleshooting, you can use separate **display** commands to collect running status data module by module, or use the **display diagnostic-information** command to bulk collect running data for multiple modules. The **display diagnostic-information** command equals this set of commands: **display clock**, **display version**, **display device**, and **display current-configuration**.

Task	Command	Remarks
Display system version information.	display version [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the system time and date.	display clock [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display or save operating statistics for multiple feature modules.	display diagnostic-information [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display alarm information.	display alarm [<i>slot slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display CPU usage statistics.	display cpu-usage [slot slot-number [cpu cpu-number]] [{ begin exclude include } regular-expression] display cpu-usage entry-number [offset] [verbose] [slot slot-number] [cpu cpu-number] [{ begin exclude include } regular-expression]	Available in any view
Display historical CPU usage statistics in a chart.	display cpu-usage history [task task-id] [slot slot-number [cpu cpu-number]] [{ begin exclude include } regular-expression]	Available in any view
Display hardware information.	display device [[slot slot-number] verbose] [{ begin exclude include } regular-expression]	Available in any view
Display the electronic label data for the device.	display device manuinfo [slot slot-number] [{ begin exclude include } regular-expression]	Available in any view
Display device temperature statistics.	display environment [slot slot-number] [{ begin exclude include } regular-expression]	Available in any view
Display the operating states of fan trays.	display fan [slot slot-number [fan-id]] [{ begin exclude include } regular-expression]	Available in any view
Display memory usage statistics.	display memory [slot slot-number [cpu cpu-number]] [{ begin exclude include } regular-expression]	Available in any view
Display power supply information.	display power [slot slot-number [power-id]] [{ begin exclude include } regular-expression]	Available in any view
Display the mode of the last reboot.	display reboot-type [slot slot-number] [{ begin exclude include } regular-expression]	Available in any view
Display RPS status information.	display rps [slot slot-number [rps-id]] [{ begin exclude include } regular-expression]	Available in any view
Display the configuration of the job configured by using the schedule job command.	display schedule job [{ begin exclude include } regular-expression]	Available in any view
Display the device reboot setting.	display schedule reboot [{ begin exclude include } regular-expression]	Available in any view
Display the configuration of jobs configured by using the job command.	display job [job-name] [{ begin exclude include } regular-expression]	Available in any view
Display the exception handling method.	display system-failure [{ begin exclude include } regular-expression]	Available in any view

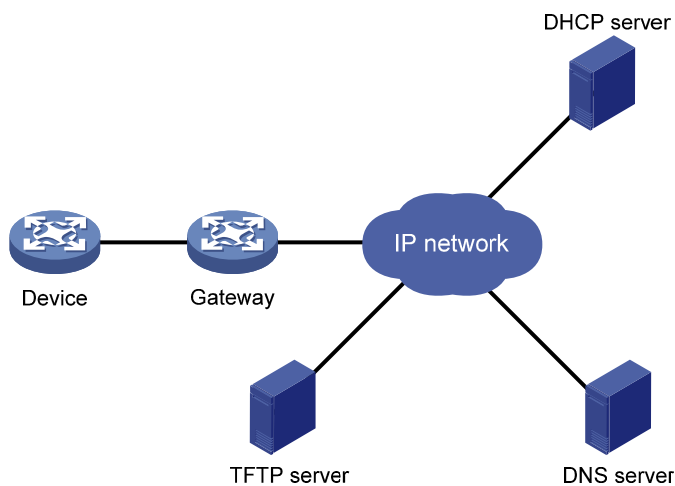
Automatic configuration introduction

Automatic configuration enables a device without any configuration file to automatically obtain and execute a configuration file during startup. Automatic configuration simplifies network configuration, facilitates centralized management, and reduces maintenance workload.

To implement automatic configuration, the network administrator saves configuration files on a server and a device automatically obtains and executes a specific configuration file.

Typical application scenario

Figure 53 Network diagram



As shown in [Figure 53](#), the device implements automatic configuration with the cooperation of the following servers:

- **DHCP server**—Assigns an IP address and other configuration parameters such as the configuration file name, TFTP server IP address, and DNS server IP address to the device.
- **TFTP server**—Saves files needed in automatic configuration. The device gets the files needed from the TFTP server, such as the host name file that saves mappings between host IP addresses and host names, and the configuration file.
- **DNS server**—Resolves between IP addresses and host names. In some cases, the device resolves its IP address to the corresponding host name through the DNS server, and then uses the host name to request the configuration file with the same name (**hostname.cfg**) from the TFTP server. If the device gets the domain name of the TFTP server from the DHCP response, the device can also resolve the domain name of the TFTP server to the IP address of the TFTP server through the DNS server.

If the DHCP server, TFTP server, DNS server, and the device are not in the same network segment, you need to configure the DHCP relay agent on the gateway, and configure routing protocols to enable each server and the device to reach one another.

How automatic configuration works

Automatic configuration works in the following manner:

1. During startup, the device sets the first up interface (if up Layer 2 Ethernet ports exist, the VLAN interface of the default VLAN of the Ethernet ports is selected as the first up interface. Otherwise, the up Layer 3 Ethernet interface with the smallest interface number is selected as the first up interface) as the DHCP client to request parameters from the DHCP server, such as an IP address and name of a TFTP server, IP address of a DNS server, and the configuration file name.
2. After getting related parameters, the device sends a TFTP request to obtain the configuration file from the specified TFTP server and executes the configuration file. If the client cannot get such parameters, it uses the factory defaults.

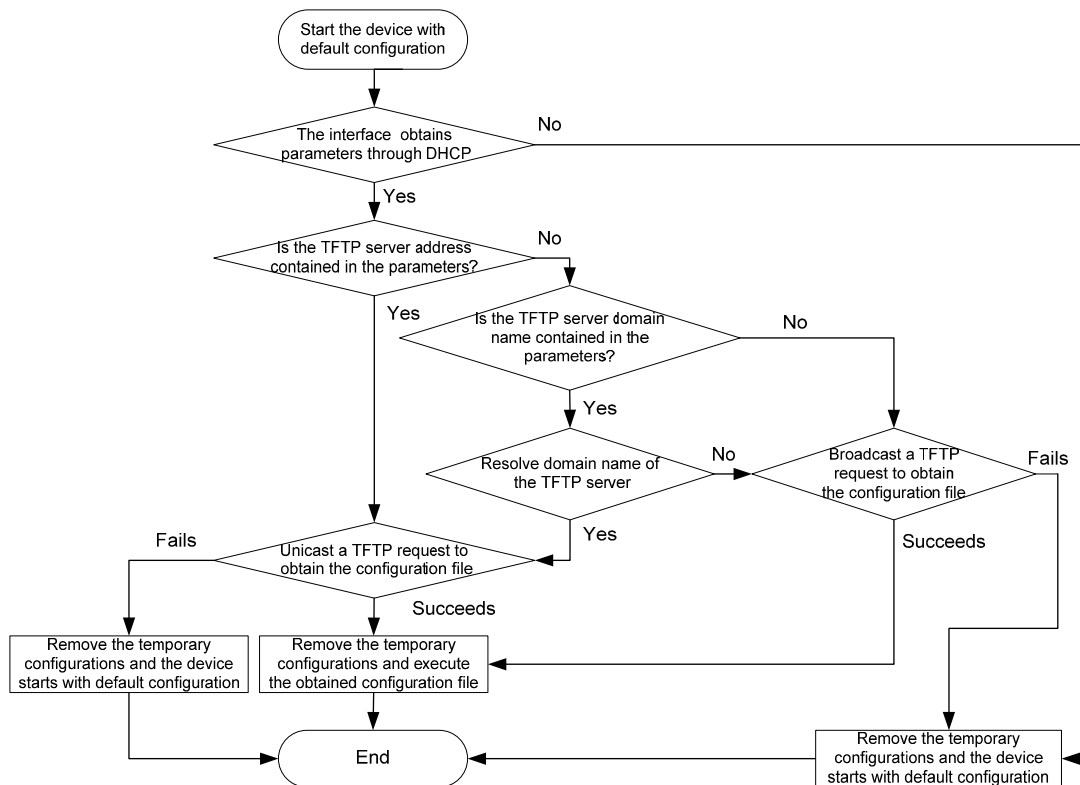
To implement automatic configuration, you must configure the DHCP server, DNS server, and TFTP server, but you do not need to perform any configuration on the device that performs automatic configuration.

Before starting the device, connect only the interface needed in automatic configuration to the network.

Automatic configuration work flow

Figure 54 shows the work flow of automatic configuration.

Figure 54 Automatic configuration work flow



Using DHCP to obtain an IP address and other configuration information

Address acquisition process

As previously mentioned, a device sets the first up interface as the DHCP client during startup. The DHCP client broadcasts a DHCP request, where the Option 55 field specifies the information that the client wants to obtain from the DHCP server such as the configuration file name, domain name and IP address of the TFTP server, and DNS server IP address.

After receiving the DHCP response from the DHCP server, the device obtains the IP address and resolves the following fields in the DHCP response:

- Option 67 or the file field—Obtains the configuration file name. The device resolves Option 67 first. If Option 67 contains the configuration file name, the device does not resolve the file field. If not, the device resolves the file field.
- Option 66—Obtains the TFTP server domain name
- Option 150—Obtains the TFTP server IP address
- Option 6—Obtains the DNS server IP address.

If no response is received from the DHCP server, the device removes the temporary configuration and starts up with the factory defaults.

The temporary configuration contains two parts: the configuration made on the interface through which automatic configuration is performed, and the **ip host** command in the host name file (For more information about the **ip host** command, see *Layer 3—IP Services Command Reference*.). The temporary configuration is removed by executing the corresponding **undo** commands.

For more information about DHCP, see *Layer 3—IP Services Configuration Guide*.

Principles for selecting an address pool on the DHCP server

The DHCP server selects IP addresses and other network configuration parameters from an address pool for clients. DHCP supports the following types of address pools:

- **Dynamic address pool**—A dynamic address pool contains a range of IP addresses and other parameters that the DHCP server dynamically assigns to clients.
- **Static address pool**—A static address pool contains the binding of an IP address and a MAC address (or a client ID). The DHCP server assigns the IP address of the binding and specific configuration parameters to a requesting client whose MAC address or ID is contained in the binding. In this way, the client can get a fixed IP address.

Select address pools by using one of the following methods:

- If devices use the same configuration file, you can configure a dynamic address pool on the DHCP server to assign IP addresses and the same configuration parameters (for example, configuration file name) to the devices. In this case, the configuration file can only contain common configurations of the devices, and the specific configurations of each device need to be performed in other ways. For example, the configuration file can enable Telnet and create a local user on devices so that the administrator can Telnet to each device to perform specific configurations (for example, configure the IP address of each interface).
- If devices use different configuration files, you need to configure static address pools to make sure each device can get a fixed IP address and a specific configuration file. With this method, no more configuration is required for the devices.

To configure static address pools, you must obtain corresponding client IDs. To obtain a device's client ID, use the **display dhcp server ip-in-use** command to display address binding information on the DHCP server after the device obtains its IP address through DHCP.

Obtaining the configuration file from the TFTP server

A device can obtain the following files from the TFTP server during automatic configuration:

- The configuration file specified by the Option 67 or file field in the DHCP response.
- The host name file named `network.cfg` that stores mappings between IP addresses and host names.

For example, the host name file can include the following:

```
ip host host1 101.101.101.101
ip host host2 101.101.101.102
ip host client1 101.101.101.103
ip host client2 101.101.101.104
```

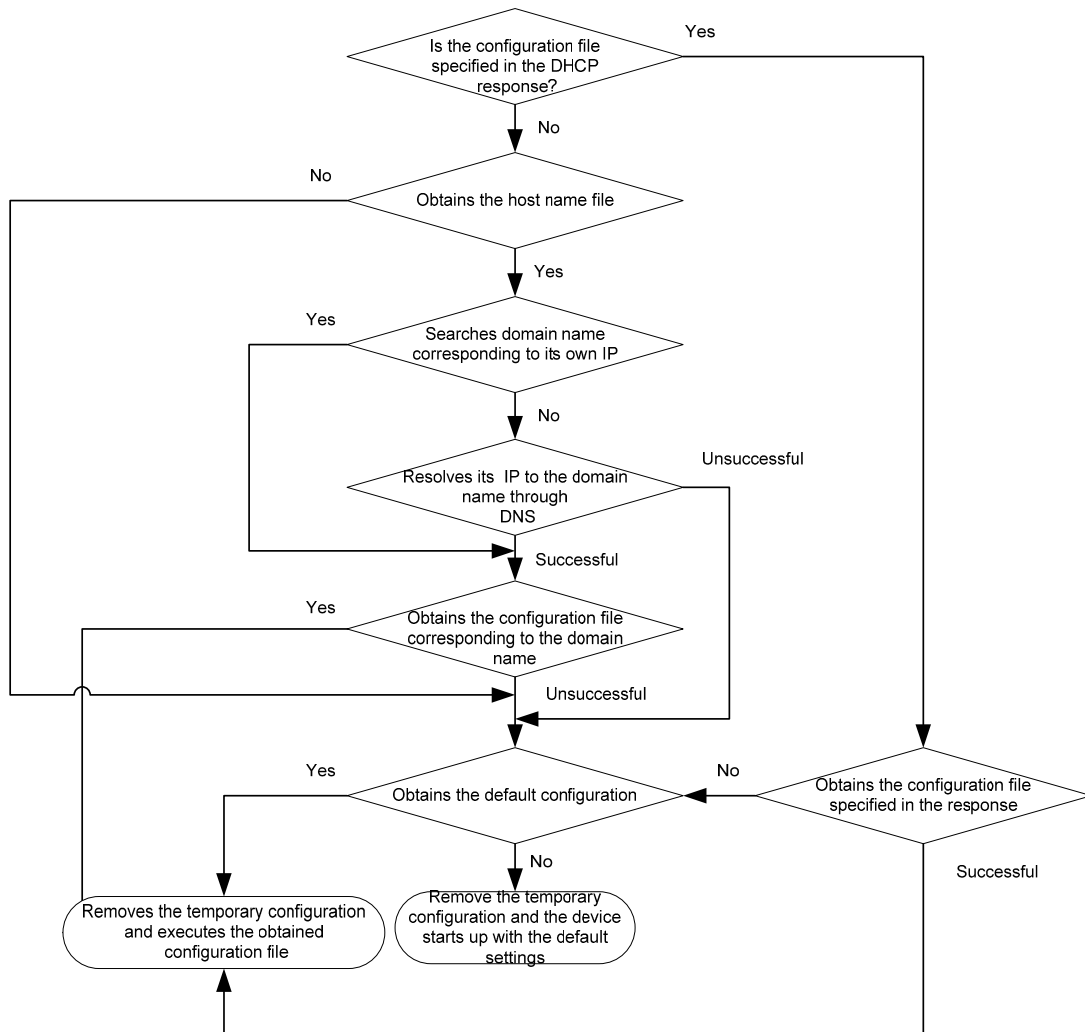


IMPORTANT:

- There must be a space before the keyword **ip host**.
 - The host name of a device saved in the host name file must be the same as the configuration file name of the device, and can be identical with or different from that saved in the DNS server.
-
- The configuration file for the device, which is named `hostname.cfg` (*hostname* is the host name of the device). For example, if the host name of a device is **aaa**, the configuration file for the device is named **aaa.cfg**.
 - The default configuration file named **device.cfg**.

Obtaining the configuration file

Figure 55 Obtaining the configuration file



A device obtains its configuration file by using the following workflow:

- If the DHCP response contains the configuration file name, the device requests the specified configuration file from the TFTP server.
- If not, the device tries to get its host name from the host name file obtained from the TFTP server. If it fails, the device resolves its IP address to the host name through DNS server. Once the device gets its host name, it requests the configuration file with the same name from the TFTP server.
- If all the above operations fail, the device requests the default configuration file from the TFTP server.

TFTP request sending mode

The device chooses whether to unicast or broadcast a TFTP request as follows:

- If a legitimate TFTP server IP address is contained in the DHCP response, the device unicasts a TFTP request to the TFTP server.
- If not, the device resolves the TFTP server domain name contained in the DHCP response to the corresponding IP address through the DNS server. If successful, the device unicasts a TFTP request to the TFTP server; if not, the device broadcasts a TFTP request.

- If the IP address and the domain name of the TFTP server are not contained in the DHCP response or they are illegitimate, the device broadcasts a TFTP request.

After broadcasting a TFTP request, the device selects the TFTP server that responds first to obtain the configuration file. If the requested configuration file does not exist on the TFTP server, the request operation fails, and the device removes the temporary configuration and starts up with the factory defaults.

If the device and the TFTP server reside in different subnets, you must configure the UDP Helper function for the gateway to change the broadcast TFTP request from the device to a unicast packet and forward the unicast packet to the specified TFTP server. For more information about UDP Helper, see *Layer 3—IP Services Configuration Guide*.

Executing the configuration file

After obtaining the configuration file, the device removes the temporary configuration and executes the configuration file. If no configuration file is obtained, the device removes the temporary configuration and starts up with the factory defaults.

NOTE:

If the configuration file contains any IRF configuration, the device does not execute the IRF configuration when executing the configuration file.

The configuration file is deleted after executed. Save the configuration by using the **save** command. Otherwise, the device has to perform automatic configuration again after reboot. For more information about the **save** command, see *Fundamentals Command Reference*.

Index

[A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [H](#) [I](#) [L](#) [M](#) [N](#) [O](#) [P](#) [R](#) [S](#) [T](#) [U](#) [V](#)

A

Accessing the CLI online help, [4](#)

B

Backing up the next-startup configuration file to a TFTP server, [96](#)

C

Changing the brand name, [134](#)

Changing the system time, [130](#)

Clearing unused 16-bit interface indexes, [142](#)

CLI views, [2](#)

Command conventions, [1](#)

Configuring banners, [135](#)

Configuring configuration rollback, [93](#)

Configuring HTTP login, [54](#)

Configuring HTTPS login, [55](#)

Configuring SNMP login, [61](#)

Configuring source IP-based SNMP login control, [67](#)

Configuring temperature thresholds for a device, [141](#)

Configuring the device name, [130](#)

Configuring the exception handling method, [137](#)

Configuring the port status detection timer, [141](#)

Configuring user privilege and command levels, [13](#)

Configuring Web login control, [69](#)

Controlling Telnet logins, [65](#)

Controlling the CLI output, [10](#)

D

Deleting the next-startup configuration file, [97](#)

Disabling Boot ROM access, [141](#)

Displaying and maintaining a configuration file, [98](#)

Displaying and maintaining CLI, [19](#)

Displaying and maintaining CLI login, [53](#)

Displaying and maintaining device management, [143](#)

Displaying and maintaining FTP, [80](#)

Displaying and maintaining software upgrade, [108](#)

Displaying and maintaining the TFTP client, [82](#)

Displaying and maintaining Web login, [57](#)

E

Enabling displaying the copyright statement, [133](#)

Entering a command, [5](#)

F

File system management examples, [88](#)

H

How automatic configuration works, [146](#)

HTTP login configuration example, [57](#)

HTTPS login configuration example, [59](#)

I

ISSU overview, [112](#)

L

Logging in through SSH, [40](#)

Logging in through Telnet, [32](#)

Logging in through the console port for the first time, [23](#)

Logging in to the CLI, [1](#)

Login methods at a glance, [21](#)

M

Managing directories, [86](#)

Managing files, [84](#)

Managing storage media, [87](#)

Modem dial-in through the console port, [43](#)

N

Network requirements, [120](#)

Network status, [120](#)

NMS login example, [63](#)

O

Overview, [90](#)

P

Performing an ISSU, [114](#)

Performing batch operations, [88](#)

Prerequisites, [81](#)

R

Rebooting the device, [137](#)

Restoring the next-startup configuration file from a TFTP server, [97](#)

S

Saving the running configuration, [91](#)

Saving the running configuration, [19](#)

Scheduling jobs, [139](#)

Setting the file system operation mode, [88](#)

Software upgrade examples, [108](#)

Software upgrade methods, [99](#)

Specifying a configuration file for the next startup, [96](#)

T

TFTP client configuration example, [82](#)

Typical application scenario, [145](#)

U

Understanding command-line error messages, [8](#)

Upgrade procedure, [121](#)

Upgrading software by installing hotfixes, [101](#)

Upgrading software through a system reboot, [100](#)

User interfaces, [22](#)

Using the command history function, [9](#)

Using the device as a TFTP client, [81](#)

Using the device as an FTP client, [71](#)

Using the device as an FTP server, [76](#)

Using the undo form of a command, [2](#)

V

Verifying and diagnosing transceiver modules, [142](#)

Contents

IRF overview	1
Hardware compatibility	1
IRF benefits	1
Application scenario	1
Basic concepts	2
IRF member roles	2
IRF member ID	2
IRF port	2
Physical IRF port	3
IRF domain ID	3
IRF split	3
IRF merge	4
Member priority	4
Interface naming conventions	4
File system naming conventions	5
Configuration synchronization mechanism	6
Master election	6
IRF multi-active detection	6
Multi-active handling procedure	6
LACP MAD	7
BFD MAD	8
ARP MAD	9
Configuring IRF	11
General restrictions and configuration guidelines	11
Software requirements	11
IRF physical port restrictions and cabling requirements	11
IRF link redundancy	11
MAD	11
Other configuration guidelines	12
Setup and configuration task list	12
Planning the IRF fabric setup	13
Assigning a member ID to each IRF member switch	14
Specifying a priority for each member switch	14
Connecting physical IRF ports	15
Connection restrictions	15
IRF fabric topology	15
Binding physical ports to IRF ports	16
Accessing the IRF fabric	17
Accessing the CLI of the master switch	17
Accessing the CLI of a subordinate switch	17
Assigning an IRF domain ID to the IRF fabric	18
Configuring a member switch description	19
Configuring IRF link load sharing mode	19
Configuring the global load sharing mode	20
Configuring port-specific load sharing criteria	20
Configuring IRF bridge MAC persistence	20
Enabling software auto-update for system software image synchronization	21
Setting the IRF link down report delay	22

Configuring MAD.....	22
Configuring LACP MAD.....	23
Configuring BFD MAD.....	24
Configuring ARP MAD.....	25
Excluding a port from the shutdown action upon detection of multi-active collision.....	26
Recovering an IRF fabric.....	27
Displaying and maintaining an IRF fabric.....	28
Configuration examples	29
LACP MAD-enabled IRF configuration example.....	29
BFD MAD-enabled IRF configuration example.....	31
ARP MAD-enabled IRF configuration example.....	34
Index	37

IRF overview

The HP Intelligent Resilient Framework (IRF) technology creates a large IRF fabric from multiple switches to provide data center class availability and scalability. IRF virtualization technology offers processing power, interaction, unified management, and uninterrupted maintenance of multiple switches.

This book describes IRF concepts and guides you through the IRF setup procedure.

Hardware compatibility

5500 switches in an IRF fabric must be the same sub-series. For example, a 5500 EI and a 5500 SI cannot form an IRF fabric.

To use IRF on these switches, you must purchase at least one 10-GE interface card listed in "[IRF physical port restrictions and cabling requirements](#)."

IRF benefits

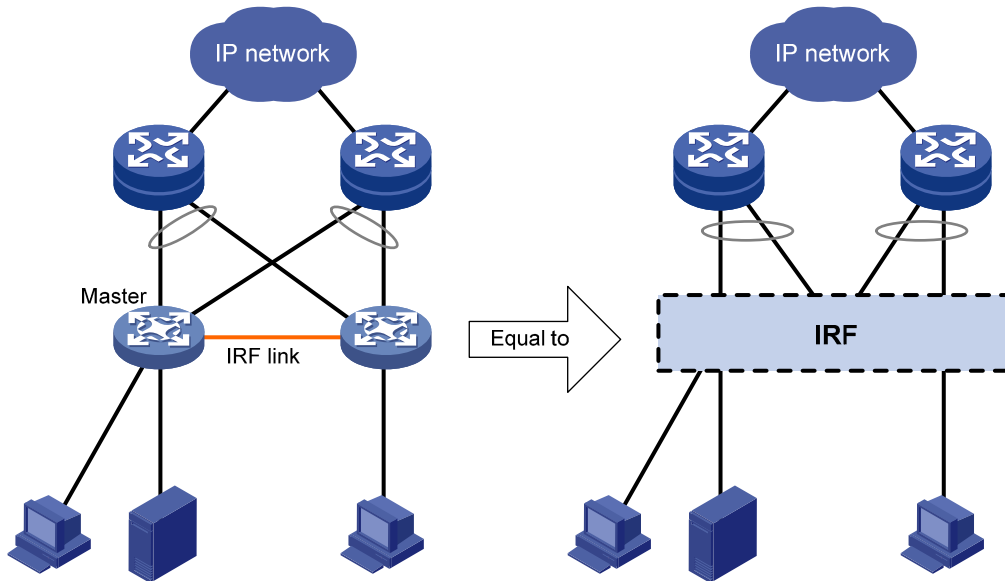
IRF delivers the following benefits:

- **Simplified topology and easy management**—An IRF fabric appears as one node and is accessible at a single IP address on the network. You can use this IP address to log in at any member device to manage all the members of the IRF fabric. In addition, you do not need to run the spanning tree feature among the IRF members.
- **1:N redundancy**—In an IRF fabric, one member works as the master to manage and control the entire IRF fabric, and all the other members process services while backing up the master. When the master fails, all the other member devices elect a new master from among them to take over without interrupting services.
- **IRF link aggregation**—You can assign several physical links between neighboring members to their IRF ports to create a load-balanced aggregate IRF connection with redundancy.
- **Multiple-chassis link aggregation**—You can use the Ethernet link aggregation feature to aggregate the physical links between the IRF fabric and its upstream or downstream devices across the IRF members.
- **Network scalability and resiliency**—Processing capacity of an IRF fabric equals the total processing capacities of all the members. You can increase ports, network bandwidth, and processing capacity of an IRF fabric simply by adding member devices without changing the network topology.

Application scenario

[Figure 1](#) shows an IRF fabric that comprises two switches, which appear as a single node to the upper and lower layer devices.

Figure 1 IRF application scenario



Basic concepts

This section describes the basic concepts that you might encounter when working with IRF.

IRF member roles

IRF uses two member roles: master and slave (called "subordinate" throughout the documentation).

When switches form an IRF fabric, they elect a master to manage the IRF fabric, and all other switches back up the master. When the master switch fails, the other switches automatically elect a new master from among them to take over. For more information about master election, see "[Master election](#)."

IRF member ID

An IRF fabric uses member IDs to uniquely identify and manage its members. This member ID information is included as the first part of interface numbers and file paths to uniquely identify interfaces and files in an IRF fabric. For more information about interface and file path naming, see "[Interface naming conventions](#)" and "[File system naming conventions](#)."

If two switches have the same IRF member ID, they cannot form an IRF fabric.

IRF port

An IRF port is a logical interface for the connection between IRF member devices. Every IRF-capable device supports two IRF ports. The IRF ports are named IRF-port $n/1$ and IRF-port $n/2$, where n is the member ID of the switch. The two IRF ports are referred to as "IRF-port 1" and "IRF-port 2" in this book for simplicity.

To use an IRF port, you must bind at least one physical port to it. The physical ports assigned to an IRF port automatically form an aggregate IRF link. An IRF port goes down only if all its physical IRF ports are down.

Physical IRF port

Physical IRF ports connect IRF member devices and must be bound to an IRF port. They forward IRF protocol packets between IRF member devices and data packets that must travel across IRF member devices.

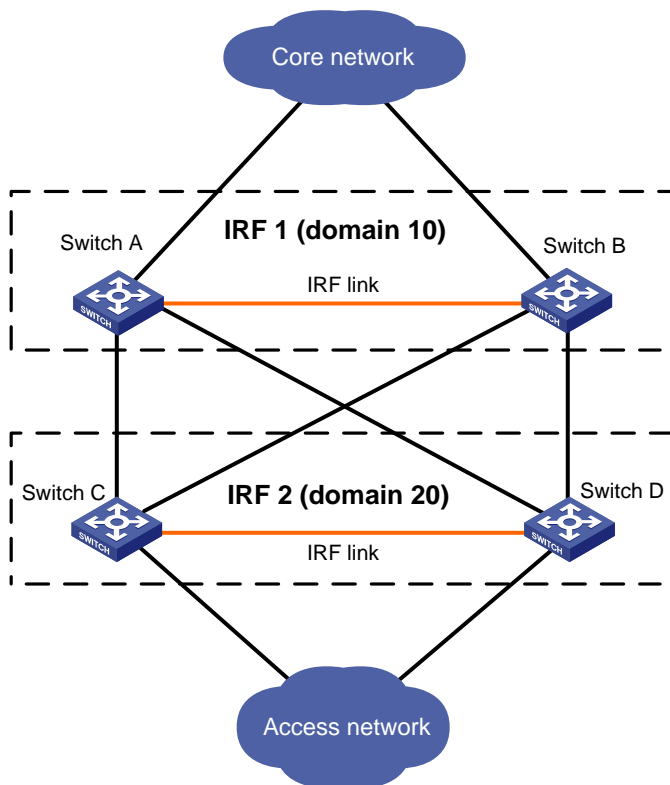
For more information about physical ports that can be used for IRF links, see "[General restrictions and configuration guidelines](#)."

IRF domain ID

One IRF fabric forms one IRF domain. IRF uses IRF domain IDs to uniquely identify IRF fabrics and prevent IRF fabrics from interfering with one another.

As shown in [Figure 2](#), Switch A and Switch B form IRF fabric 1, and Switch C and Switch D form IRF fabric 2. The fabrics have LACP MAD detection links between them. When a member switch in one IRF fabric receives an extended LACP packet for MAD detection, it looks at the domain ID in the packet to see whether the packet is from the local IRF fabric or from a different IRF fabric. Then, the switch can handle the packet correctly.

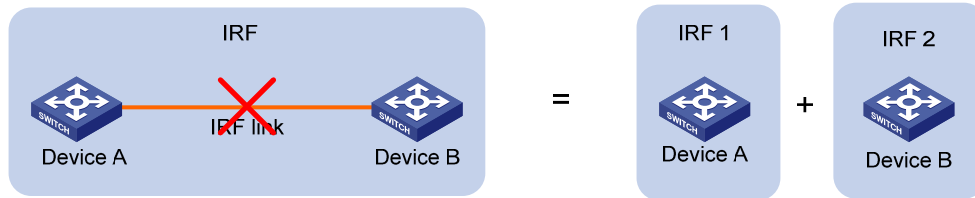
Figure 2 A network that comprises two IRF domains



IRF split

IRF split occurs when an IRF fabric breaks up into two or more IRF fabrics because of IRF link failures, as shown in [Figure 3](#). The split IRF fabrics operate with the same IP address and cause routing and forwarding problems on the network. To quickly detect a multi-active collision, configure at least one MAD mechanisms (see "[IRF multi-active detection](#)").

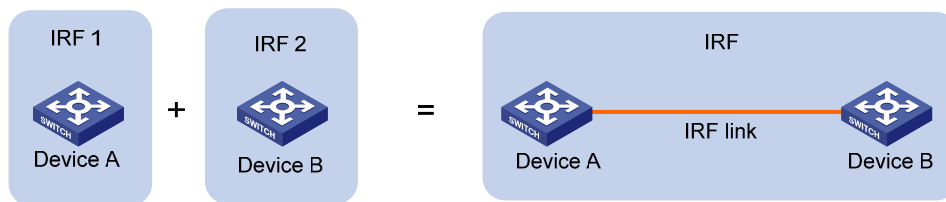
Figure 3 IRF split



IRF merge

IRF merge occurs when two split IRF fabrics re-unite or when you configure and connect two independent IRF fabrics to be one IRF fabric, as shown in Figure 4.

Figure 4 IRF merge



Member priority

Member priority determines the possibility of a member device to be elected the master. A member with higher priority is more likely to be elected the master.

The default member priority is 1. You can change the member priority of a member device to affect the master election result.

Interface naming conventions

An interface is named in the format of *member-id/slot-number/port-index*, where:

- *member-id*—If the switch is standalone, the member ID defaults to 1. If the standalone switch was once an IRF member switch, it uses the same member ID as it was in the IRF fabric.
- *slot-number*—Represents the slot number of the interface card. This argument takes 0 for the fixed ports on the front panel, and takes 1 and 2 for the two interface cards (from left to right) on the rear panel.
- *port-index*—Port index depends on the number of ports available on the switch. To identify the index of a port, look at its port index mark on the chassis.

For one example, on the standalone switch Sysname, GigabitEthernet 1/0/1 represents the first fixed port on the front panel. Set its link type to trunk, as follows:

```
<Sysname> system-view
[Sysname] interface gigabitethernet 1/0/1
[Sysname-GigabitEthernet1/0/1] port link-type trunk
```

For another example, on the IRF fabric Master, GigabitEthernet 3/0/1 represents the first fixed port on the front panel of member switch 3. Set its link type to trunk, as follows:

```
<Sysname> system-view
```

```
[Sysname] interface gigabitethernet 3/0/1
[Sysname-GigabitEthernet3/0/1] port link-type trunk
```

File system naming conventions

On a standalone switch, you can use the name of storage device to access its file system. For more information about storage device naming conventions, see *Fundamentals Configuration Guide*.

On an IRF fabric, you can use the name of storage device to access the file system of the master. To access the file system of any other member switch, use the name in the **slotmember-ID#storage-device-name** format. For example:

To access the **test** folder under the root directory of the Flash on the master switch:

```
<Master> mkdir test
...
%Created dir flash:/test.
<Master> dir
Directory of flash:/
  0   -rw-  10105088  Apr 26 2000 13:44:57  test.bin
  1   -rw-    2445    Apr 26 2000 15:18:19  config.cfg
  2   drw-      -    Jul 14 2008 15:20:35  test
30861 KB total (20961 KB free)
```

To create and access the **test** folder under the root directory of the Flash on member switch 3:

```
<Master> mkdir slot3#flash:/test
%Created dir slot3#flash:/test.
<Master> cd slot3#flash:/test
<Master> pwd
slot3#flash:/test
```

Or:

```
<Master> cd slot3#flash:/
<Master> mkdir test
%Created dir slot3#flash:/test.
```

To copy the file **test.bin** on the master to the root directory of the Flash on member switch 3:

Display the current working path. In this example, the current working path is the root directory of the Flash on member switch 3.

```
<Master> pwd
slot3#flash:
```

Change the current working path to the root directory of the Flash on the master switch:

```
<Master> cd flash:/
<Master> pwd
flash:
```

Copy the file to member switch 3.

```
<Master> copy test.bin slot3#flash:/
Copy flash:/test.bin to slot3#flash:/test.bin?[Y/N]:y
%Copy file flash:/test.bin to slot3#flash:/test.bin...Done.
```

Configuration synchronization mechanism

IRF uses a strict running-configuration synchronization mechanism so all chassis in an IRF fabric can work as a single node, and after the master fails, other members can operate normally.

In an IRF fabric, all chassis get and run the running configuration of the master. Any configuration you have made is propagated to all members.

When you execute the **save [safely] [backup | main] [force]** command or the **save file-url all** command, the system saves the running configuration, as follows:

- If the configuration auto-update function (the **slave auto-update config** command) is enabled, saves the configuration as the startup configuration on all member switches for the next startup.
- If the configuration auto-update function is disabled, saves the configuration as the startup configuration on the master for the next startup.

For more information about configuration management, see *Fundamentals Configuration Guide*.

Master election

Master election is held each time the IRF fabric topology changes, for example, when the IRF fabric is established, a new member device is plugged in, the master device fails or is removed, the IRF fabric splits, or IRF fabrics merge.

Master election uses the following rules in descending order:

1. Current master, even if a new member has higher priority.
When an IRF fabric is being formed, all member switches consider themselves as the master, and this rule is skipped
2. Member with higher priority.
3. Member with the longest system uptime.
4. Member with the lowest bridge MAC address.

The IRF fabric is formed on election of the master.

During an IRF merge, the switches of the IRF fabric that fails the master election reboot automatically to re-join the IRF fabric that wins the election.

After a master election, all subordinate switches initialize and reboot with the configuration on the master. Their original configuration, even if has been saved, does not take effect.

IRF multi-active detection

An IRF link failure causes an IRF fabric to split in two IRF fabrics operating with the same Layer 3 configurations, including the same IP address. To avoid IP address collision and network problems, IRF uses multi-active detection (MAD) mechanisms to detect the presence of multiple identical IRF fabrics, handle collisions, and recover from faults.

Multi-active handling procedure

The multi-active handling procedure includes detection, collision handling and failure recovery.

Detection

The MAD implementation of this switch series detects active IRF fabrics with the same Layer 3 global configuration by extending the LACP, BFD, or gratuitous ARP protocol.

These MAD mechanisms identify each IRF fabric with a domain ID and an active ID (the member ID of the master). If multiple active IDs are detected in a domain, MAD determines that an IRF collision or split has occurred.

You can use at least one of these mechanisms in an IRF fabric, depending on your network topology. For a comparison of these MAD mechanisms, see "[Configuring MAD](#)."

Collision handling

When multiple identical active IRF fabrics are detected, MAD compares the member IDs of their masters. If the master in one IRF fabric has the lowest member ID among all the masters, the members in the fabric continue to operate in Active state and forward traffic. MAD sets all the other IRF fabrics in Recovery (disabled) state and shuts down all their physical ports but the console ports, physical IRF ports, and any ports you have specified with the **mad exclude interface** command.

Failure recovery

To merge two split IRF fabrics, first repair the failed IRF link and remove the IRF link failure.

If the IRF fabric in Recovery state fails before the failure is recovered, repair the failed IRF fabric and the failed IRF link.

If the IRF fabric in Active state fails before the failure is recovered, first enable the IRF fabric in Recovery state to take over the active IRF fabric and protect the services from being affected. After that, recover the MAD failure.

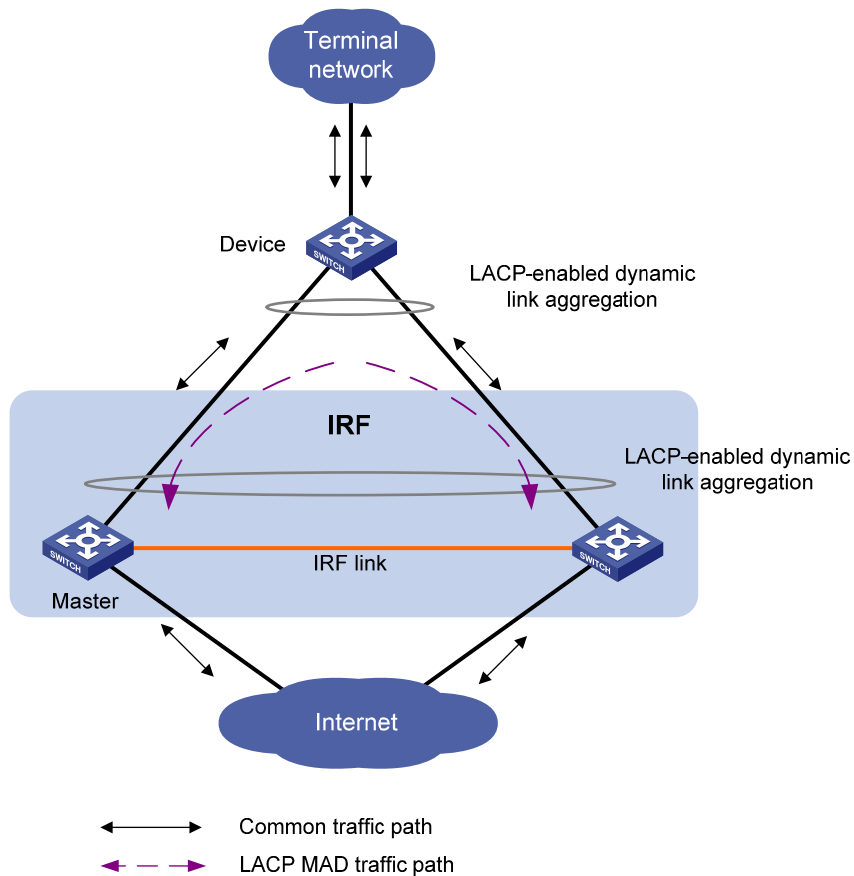
LACP MAD

LACP MAD requires that every IRF member have a link with an intermediate device, and all these links form a dynamic link aggregation group, as shown in [Figure 5](#). In addition, the intermediate device must be an HP device that supports extended LACP for MAD.

The IRF member switches send extended LACPDU s with TLVs that convey the domain ID and the active ID of the IRF fabric. The intermediate device transparently forwards the extended LACPDU s received from one member switch to all the other member switches:

- If the domain IDs and the active IDs in the extended LACPDU s sent by all the member devices are the same, the IRF fabric is integrated.
- If the extended LACPDU s convey the same domain ID but different active IDs, a split has occurred. To handle this situation, LACP MAD sets the IRF fabric with higher active ID in Recovery state, and shuts down all its physical ports but the console port, IRF ports, and any ports you have specified with the **mad exclude interface** command. The IRF fabric with lower active ID is still in Active state and forwards traffic.

Figure 5 LACP MAD application scenario



BFD MAD

BFD MAD can work with or without intermediate devices. Figure 6 shows a typical BFD MAD application scenario.

To use BFD MAD:

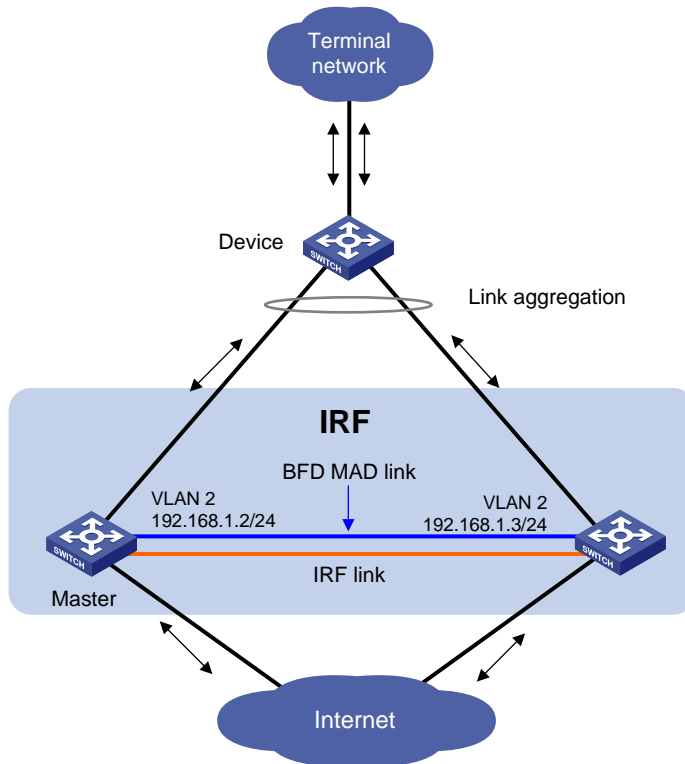
- Set up dedicated BFD MAD link between each pair of IRF members or between each IRF member and the intermediate device. Do not use the BFD MAD links for any other purpose.
- Assign the ports connected by BFD MAD links to the same VLAN, create a VLAN interface for the VLAN, and assign a MAD IP address to each member on the VLAN interface.

The MAD addresses identify the member switches and must belong to the same subnet.

With BFD MAD, the master tries to establish BFD sessions with other member switches by using its MAD IP address as the source IP address:

- If the IRF fabric is integrated, only the MAD IP address of the master is effective, and the master cannot establish a BFD session with any other member. If you execute the **display bfd session** command, the state of the BFD sessions is **Down**.
- When the IRF fabric splits, the IP addresses of the masters in the split IRF fabrics take effect, and the two masters can establish a BFD session. If you execute the **display bfd session** command, the state of the BFD session between the two devices is **Up**.

Figure 6 BFD MAD application scenario

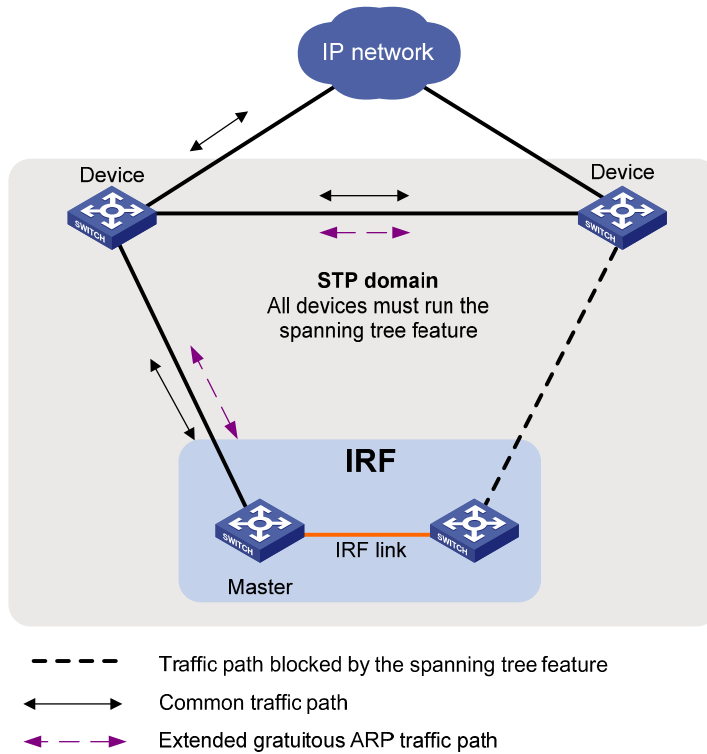


ARP MAD

ARP MAD detects multi-active collisions by using extended gratuitous ARP packets that convey the IRF domain ID and the active ID.

You can set up ARP MAD links between neighbor IRF member devices, or more commonly, between each IRF member device and an intermediate device (see [Figure 7](#)). If an intermediate device is used, you must also run the spanning tree feature between the IRF fabric and the intermediate device.

Figure 7 ARP MAD application scenario



Each IRF member compares the domain ID and the active ID in incoming extended gratuitous ARP packets with its domain ID and active ID:

- If the domain IDs are different, the extended gratuitous ARP packet is from a different IRF fabric, and the device does not continue to process the packet with the MAD mechanism.
- If the domain IDs are the same, the device compares the active IDs:
 - If the active IDs are different, the IRF fabric has split.
 - If the active IDs are the same, the IRF fabric is integrated.

Configuring IRF

Read the configuration restrictions and guidelines carefully when you connect and set up an IRF fabric.

General restrictions and configuration guidelines

This section describes the restrictions and configuration guidelines you must follow.

Software requirements

All IRF member switches must run the same system software image version.

IRF physical port restrictions and cabling requirements

Candidate IRF physical ports include 10-GE ports on expansion interface cards. To use IRF, purchase at least one interface card. The following are the interface cards available for IRF connections:

- LSPM1XP1P (JD361B)
- LSPM1XP2P (JD359B)
- LSPM1CX2P (JD360B)
- LSPM2SP2P (JD368B)

For long-distance IRF connections, use XFP or SFP+ transceiver modules and fibers. For short-distance IRF connections, use CX4 or SFP+ cables. For more information about transceiver modules available for an interface card, see the interface card user guide.

For more information about transceiver modules, see *HP A-Series Switches Transceiver Modules User Guide*.

The SFP+ modules and SFP+ cables available for this switch series are subject to change over time. For the most up-to-date list of SFP+ modules and cables, contact HP technical support or marketing staff.

IRF link redundancy

This switch series supports up to two physical ports to one IRF port. These two ports must be located on the same interface card.

MAD

- Configure at least one MAD mechanism for prompt IRF split detection and IRF fabric recovery.
- If LACP MAD or ARP MAD runs between two IRF fabrics, assign each fabric a unique IRF domain ID. For BFD MAD, this task is optional.
- BFD MAD is available only on 5500 EI switches.
- To exclude a port from the shutdown action performed when an IRF fabric to the Recovery state, use the **mad exclude interface** command. To bring up a port after the IRF fabric transits to the Recovery state, you must use the **mad restore** command to activate the entire IRF fabric, rather than using the **undo shutdown** command.

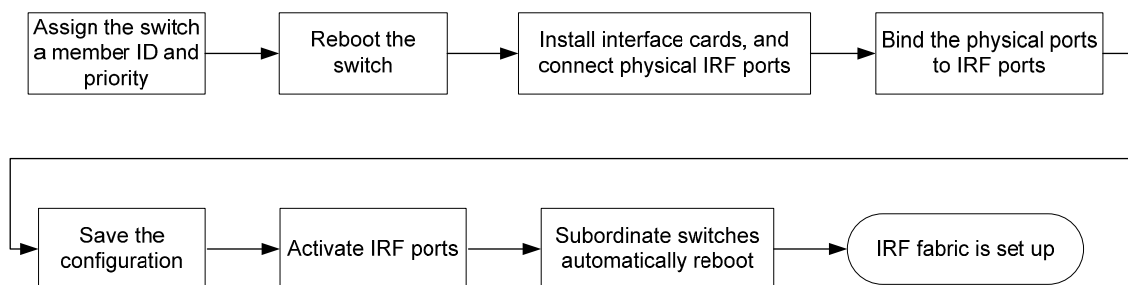
Other configuration guidelines

- Strictly follow the IRF fabric setup procedure described in "[Setup and configuration task list](#)" to plan the IRF fabric, identify IRF physical ports, connect IRF member switches, and configure basic settings.
- Assign each member a unique IRF member ID to make sure they can merge. You must reboot the members to validate the IRF member ID settings.
- Assign the highest member priority to the device you want to use as the master.
- Before removing an interface card that has physical IRF ports in an IRF fabric, remove the IRF connection cables, or use the **shutdown** command to shut down the IRF physical ports.
- If a subordinate switch uses the same next-startup configuration file name as the master switch, the file might be overwritten depending on your configuration file management settings. To continue to use the configuration file after removing the switch from the IRF fabric, back up the file before setting up the IRF fabric.
- Save any configuration you have made to the startup configuration file before rebooting the IRF member devices.
- The Layer 3 Ethernet port in this book refers to an Ethernet port that can perform IP routing and inter-VLAN routing. You can set an Ethernet port as a Layer 3 Ethernet interface by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).
- You can configure an Ethernet port as a Layer 3 Ethernet port only on the 5500 EI switch.

Setup and configuration task list

HP recommends the configuration procedure in [Figure 8](#).

Figure 8 IRF setup and configuration flow chart



HP recommends the following IRF fabric setup and configuration procedure:

Task	Remarks
1. Planning the IRF fabric setup	Required.
2. Assigning a member ID to each IRF member switch	Required. Perform this task on each member switch.
3. Specifying a priority for each member switch	Required. Perform this task on each member switch.
4. Connecting physical IRF ports	Required.

Task	Remarks
5. Binding physical ports to IRF ports	Required. Perform this task on each member switch.
6. Accessing the IRF fabric: <ul style="list-style-type: none"> o Accessing the CLI of the master switch o Accessing the CLI of a subordinate switch 	Login to the master's CLI is required. You configure all member switches at the master's CLI. From the master's CLI, you can log in to any other member switch's CLI to execute a limited set of maintenance commands.
7. Assigning an IRF domain ID to the IRF fabric	This task is required for ARP MAD and LACP MAD.
8. Configuring a member switch description	Optional.
9. Configuring IRF link load sharing mode: <ul style="list-style-type: none"> o Configuring the global load sharing mode o Configuring port-specific load sharing criteria 	Optional.
10. Configuring IRF bridge MAC persistence	Optional.
11. Enabling software auto-update for system software image synchronization	Optional. HP recommends enabling software auto-update to make sure system software image synchronization
12. Setting the IRF link down report delay	Optional.
13. Configuring MAD: <ul style="list-style-type: none"> o Configuring LACP MAD o Configuring BFD MAD o Configuring ARP MAD o Excluding a port from the shutdown action upon detection of multi-active collision o Recovering an IRF fabric 	Required. MAD mechanisms are independent of one another. You can configure at least one MAD mechanism for an IRF fabric.

Planning the IRF fabric setup

Consider the following items when you plan an IRF fabric:

- Hardware compatibility and restrictions
- IRF fabric size
- Master switch
- IRF physical ports
- Member ID and priority assignment scheme
- Fabric topology and cabling scheme

For more information about hardware and cabling, see the installation guide for the device.

Assigning a member ID to each IRF member switch

CAUTION:

In an IRF fabric, changing IRF member IDs might cause undesirable configuration changes and even data loss. Before you do that, back up the configuration and make sure you fully understand the impact on your network. For example, all member switches in an IRF fabric are the same model. If you swapped the IDs of any two members, their interface settings would also be swapped.

By default, the member IDs of all switches are 1. To create an IRF fabric, you must assign a unique IRF member ID to each switch.

Perform this task before the IRF fabric is formed. To prevent any undesirable configuration change or data loss, avoid changing member IDs after the IRF fabric is formed.

The new member ID takes effect at a reboot. After the switch reboots, the settings on all member-ID related physical resources (including common physical network ports) are removed and require reconfiguration, regardless of whether you have saved the configuration.

To set a member ID for a switch:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Assign an IRF member ID to the switch.	irf member <i>member-id</i> renumber <i>new-member-id</i>	The default IRF member ID is 1. Optional.
3. Save the configuration.	save [safely] [backup main] [force]	If you have bound physical ports to IRF ports or assigned member priority, save the configuration before rebooting the switch so these settings can continue to take effect after the reboot.
4. Reboot the switch.	reboot [slot <i>slot-number</i>]	N/A

Specifying a priority for each member switch

IRF member priority represents the possibility for a device to be elected the master in an IRF fabric. The higher the priority, the higher the possibility.

A member priority change affects the election result at the next master election, but does not cause immediate master re-election.

To specify a priority for the switch:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Specify a priority for the switch.	irf priority <i>priority</i>	The default IRF member priority is 1.

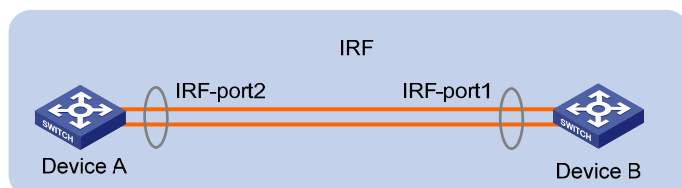
Connecting physical IRF ports

When you connect two neighboring IRF members, connect the physical ports of IRF-port 1 on one member to the physical ports of IRF-port 2 on the other, as shown in [Figure 9](#).

! **IMPORTANT:**

No intermediate devices are allowed between neighboring members.

Figure 9 Connecting IRF physical ports



Connection restrictions

The switches support up to two physical IRF ports for an IRF port.

Note the following IRF link binding restrictions when you connect physical IRF ports:

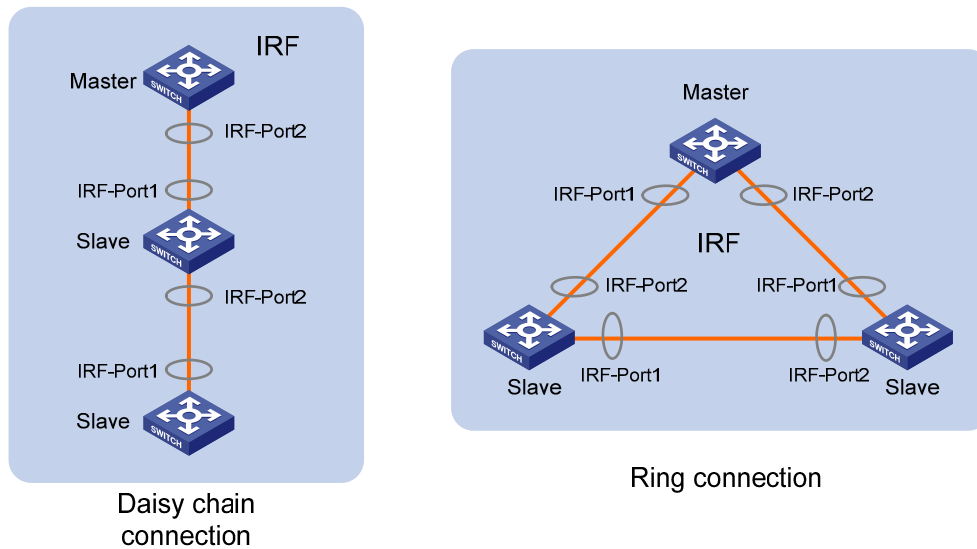
- Facing the rear panel, make sure the physical ports bound to IRF port 1 is on the left of the physical ports bound to IRF port 2.
- Physical ports of an IRF port must be on the same interface card

IRF fabric topology

Connect the switches into a daisy chain topology or more reliably, a ring topology (see [Figure 10](#)). In ring topology, the failure of one IRF link does not cause the IRF fabric to split as in daisy chain topology. Rather, the IRF fabric changes to a daisy chain topology without interrupting network services.

To use the ring topology, you must have at least three member switches.

Figure 10 Daisy chain topology vs. ring topology



Binding physical ports to IRF ports

When you bind physical ports to IRF ports, follow the restrictions in "[Connection restrictions](#)."

On a physical port that has been bound to an IRF port, you can only use the **bfd**, **default**, **shutdown**, **description**, and **flow-interval** commands. For more information about these commands, see *Layer 2—LAN Switching Command Reference*.

To bind physical ports to IRF ports:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter physical IRF port view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Shut down the port.	shutdown	Always shut down a physical port before binding it to an IRF port or removing the binding. Start the shutdown operation on the master and then the switch that has the fewest number of hops from the master.
4. Return to system view.	quit	N/A
5. Create an IRF port and enter IRF port view.	irf-port <i>member-id/port-number</i>	N/A
6. Bind the physical port to the IRF port.	port group interface <i>interface-type</i> <i>interface-number</i> [mode { enhanced normal }]	By default, no physical port is bound to any IRF port. The switch does not support the enhanced keyword. Make sure the two ends of an IRF link use the same binding mode.

Step	Command	Remarks
7. Return to system view.	quit	N/A
8. Enter physical IRF port view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
9. Bring up the port.	undo shutdown	N/A
10. Return to system view.	quit	N/A
11. Save the running configuration.	save	N/A
12. Activate the IRF port configuration.	irf-port-configuration active	<p>After this step is performed, the state of the IRF port changes to UP, the member switches automatically elects a master, and the subordinate switch automatically reboots.</p> <p>After the IRF fabric is formed, you can add more physical ports to an IRF port (in UP state) without performing this step.</p>

Accessing the IRF fabric

The IRF fabric appears as one device after it is formed. You configure and manage all IRF members at the CLI of the master. All settings you made are automatically propagated to the IRF members.

When you log in to an IRF fabric, you are placed at the CLI of the master, regardless of at which member switch you are logged in. After that, you can access the CLI of a subordinate switch to execute a limited set of maintenance commands.

The IRF fabric supports up to 16 concurrent VTY users. The maximum number of concurrent console users equals the total number of member switches in the IRF fabric.

Accessing the CLI of the master switch

Access an IRF fabric in one of the following ways:

- **Local login**—Log in through the console port of any member switch.
- **Remote login**—Remotely log in at a Layer 3 interface on any member switch through Telnet, SNMP, or Web.

For more information, see the chapter on login in *Fundamentals Configuration Guide*.

Accessing the CLI of a subordinate switch

You can log in to the CLI of a subordinate switch for maintenance or debugging. At the CLI of a subordinate switch, you are placed in user view, and the command prompt changes to `<Sysname-Slave#member-ID/slot-number>`, for example, `<Sysname-Slave#2>`. You can use the following commands at a subordinate switch's CLI:

- **display**
- **quit**
- **return**

- **system-view**
- **debugging**
- **terminal debugging**
- **terminal logging**
- **terminal monitor**
- **terminal trapping**

Perform the following task in user view:

Task	Command	Remarks
Log in to a subordinate switch.	irf switch-to <i>member-id</i>	By default, you are placed at the master's CLI.

To return to the master's CLI, use the **quit** command.

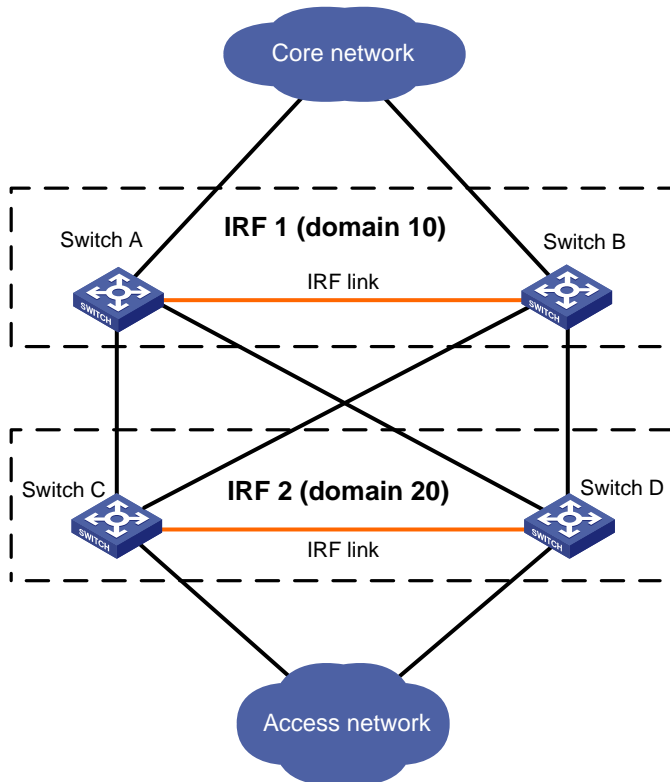
Assigning an IRF domain ID to the IRF fabric

This task is required for running LACP MAD or ARP MAD between two IRF fabrics. For BFD MAD, this task is optional.

One IRF fabric forms one IRF domain. IRF domain IDs prevent IRF fabrics from interfering with one another.

In [Figure 11](#), Switch A and Switch B form IRF fabric 1, and Switch C and Switch D form IRF fabric 2. These fabrics have LACP MAD links between them. When a member switch in one IRF fabric receives an LACP MAD packet, it looks at the domain ID in the packet to see whether the packet is from the local IRF fabric or from a different IRF fabric. Then, the switch can handle the packet correctly.

Figure 11 A network that comprises two IRF domains



To assign a domain ID to an IRF fabric:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Assign a domain ID to the IRF fabric.	irf domain <i>domain-id</i>	By default, the domain ID of an IRF fabric is 0.

Configuring a member switch description

You can configure a description for a member switch to identify its physical location, or for any other management purpose.

To configure a description for a member switch:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the description of a member.	irf member <i>member-id</i> description <i>text</i>	By default, no member switch description is configured.

Configuring IRF link load sharing mode

On an IRF port that has multiple links, traffic is balanced across its physical links. You can configure the IRF port to distribute traffic based on certain criteria, including source IP address, destination IP address,

source MAC address, destination MAC address, or a combination of them. If a criteria combination is not supported, the system displays an error message.

Configure the IRF link load sharing mode for IRF links in system view or IRF port view.

- In system view, the configuration is global and takes effect on all IRF ports.
- In IRF port view, the configuration is port specific and takes effect only on the specific IRF port.

An IRF port preferentially uses the port-specific load sharing mode. If no port-specific load sharing mode is available, it uses the global load sharing mode.

Configuring the global load sharing mode

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the global IRF link load sharing mode.	irf-port load-sharing mode { destination-ip destination-mac source-ip source-mac } *	By default, the switch uses source and destination MAC addresses for packets that have no IP header, and uses source and destination IP addresses for IP packets.

Configuring port-specific load sharing criteria

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IRF port view.	irf-port <i>member-id/port-number</i>	N/A
3. Configure the port-specific load sharing mode.	irf-port load-sharing mode { destination-ip destination-mac source-ip source-mac } *	By default, the switch uses source and destination MAC addresses for packets that have no IP header, and uses source and destination IP addresses for IP packets.

Configuring IRF bridge MAC persistence

An IRF fabric by default uses the bridge MAC address of the master switch as its bridge MAC address. This bridge MAC address is used by Layer 2 protocols, for example, LACP, to identify the IRF fabric, and must be unique on a switched LAN for proper communication.

To avoid duplicate bridge MAC addresses, an IRF fabric can automatically change its bridge MAC address after its master leaves, but the change can cause transient traffic interruption.

Depending on your network condition, enable the IRF fabric to preserve or change its bridge MAC address after the master leaves. Available options include:

- **irf mac-address persistent timer**—Bridge MAC address of the IRF fabric persists for six minutes after the master leaves. If the master does not come back before the timer expires, the IRF fabric uses the bridge MAC address of the new master as its bridge MAC address. This option avoids unnecessary

bridge MAC address change due to a device reboot, transient link failure, or purposeful link disconnection.

- **irf mac-address persistent always**—Bridge MAC address of the IRF fabric does not change after the master leaves.
- **undo irf mac-address persistent**—Bridge MAC address of the new master replaces the original one as soon as the old master leaves.

IMPORTANT:

If ARP MAD is used, configure the **undo irf mac-address persistent** command to enable immediate bridge MAC address change after a master leaves.

If two IRF fabrics have the same bridge MAC address, they cannot merge.

To configure the IRF bridge MAC persistence setting:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure IRF bridge MAC persistence.	<ul style="list-style-type: none">• Keep the bridge MAC address unchanged even if the master has changed: irf mac-address persistent always• Preserve the bridge MAC address for six minutes after the master leaves: irf mac-address persistent timer• Change the bridge MAC address as soon as the master leaves: undo irf mac-address persistent	By default, the IRF fabric's bridge MAC address persists permanently even after the master leaves.

Enabling software auto-update for system software image synchronization

To join an IRF fabric, a switch must use the same system software image as the master in the fabric.

The software auto-update function automatically propagates the system software image of the master to all members in the IRF fabric. If software auto-update is disabled, you must manually update the switch with the system software image of the master.

When you add a switch to the IRF fabric, the software auto-update function compares the system software versions of the switch and the IRF master. If the versions are different, the switch automatically downloads the system software image from the master, sets the downloaded image as the system software for the next startup, and automatically reboots with the new system software image to re-join the IRF fabric.

Before you use the software auto-update function, make sure:

- The switch you are adding to the IRF fabric is compatible with the software version running on the master. If not, the automatic system software upgrading function cannot correctly work.
- The switch you are adding to the IRF fabric has sufficient space for the new system software image.

To enable the IRF fabric to automatically synchronize the system software of the master to the switch you are adding to the IRF fabric:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable the software auto-update function.	irf auto-update enable	By default, this function is disabled.

In an IRF fabric enabled with software auto-update, if a software upgrade requires upgrading the Boot ROM image, use the following upgrading procedure:

1. Download the new system software image to the master device.
2. Use the **bootrom update** command to upgrade the Boot ROM image on the master.
This step guarantees that the master can complete startup prior to other member switches.
3. Use the **boot-loader file** *file-url slot slot-number main* command to specify the system software image as the startup image for the master.
4. Reboot the entire IRF fabric to complete upgrading software.

For the system software image and Boot ROM compatibility, see the release notes for the new software release.

Setting the IRF link down report delay

You can avoid IRF link flapping causing frequent IRF splits and merges during a short time by configuring the IRF ports to delay reporting link down events. An IRF port works as follows:

- When the IRF link changes from up to down, the port does not immediately report the change to the IRF fabric. If the IRF link state is still down when the delay time is reached, the port reports the change to the IRF fabric.
- When the IRF link changes from down to up, the link layer immediately reports the event to the IRF fabric.

To set the IRF link down report delay:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the IRF link down report delay.	irf link-delay <i>interval</i>	The default IRF link down report delay is 4 seconds. The recommended value range is 200 to 500 milliseconds. The greater the interval, the slower the service recovery.

Configuring MAD

You have the following MAD mechanisms for detecting multi-active collisions in different network scenarios:

- LACP MAD
- BFD MAD
- ARP MAD

These MAD detection mechanisms operate independently. You can configure all of them for an IRF fabric.

Table 1 provides a reference for you to make a MAD mechanism selection decision.

Table 1 A comparison of the MAD mechanisms

MAD mechanism	Advantages	Disadvantages	Application scenario
LACP MAD	<ul style="list-style-type: none"> Detection speed is fast. Requires no MAD-dedicated physical ports or interfaces. 	Requires an intermediate HP device that supports LACP MAD packets.	<p>Link aggregation is used between the IRF fabric and its upstream or downstream device.</p> <p>For information about LACP, see <i>Layer 2—LAN Switching Configuration Guide</i>.</p>
BFD MAD	<ul style="list-style-type: none"> Detection speed is fast. No intermediate device is required. Intermediate device, if used, can come from any vendor. 	<ul style="list-style-type: none"> Requires MAD dedicated physical ports and Layer 3 interfaces, which cannot be used for transmitting user traffic. If no intermediate device is used, the IRF members must be fully meshed. If an intermediate device is used, every IRF member must connect to the intermediate device. 	<ul style="list-style-type: none"> Suitable for various network scenarios. If no intermediate device is used, this mechanism is only suitable for IRF fabrics that have a small number of members that are geographically close to one another. <p>For information about BFD, see <i>High Availability Configuration Guide</i>.</p>
ARP MAD	<ul style="list-style-type: none"> No intermediate device is required. Intermediate device, if used, can come from any vendor. Requires no MAD dedicated ports. 	<ul style="list-style-type: none"> Detection speed is slower than BFD MAD and LACP MAD. MSTP must be enabled. 	<p>MSTP-enabled non-link aggregation IPv4 network scenario.</p> <p>For information about ARP, see <i>Layer 3—IP Services Configuration Guide</i>.</p>

Configuring LACP MAD

When you use LACP MAD, follow these guidelines:

- The intermediate device must be an HP device that support extended LACP for MAD.
- If the intermediate device is in an IRF fabric, assign this fabric a different domain ID than the LACP MAD-enabled fabric to avoid false detection of IRF partition.
- Use dynamic link aggregation mode. MAD is LACP dependent. Even though LACP MAD can be configured on both static and dynamic aggregate interfaces, it takes effect only on dynamic aggregate interfaces.
- Configure link aggregation settings also on the intermediate device.

To configure LACP MAD:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Assign a domain ID to the IRF fabric.	irf domain <i>domain-id</i>	The default IRF domain ID is 0.
3. Create an aggregate interface and enter aggregate interface view.	<ul style="list-style-type: none"> Enter Layer 2 aggregate interface view: interface bridge-aggregation <i>interface-number</i> Enter Layer 3 aggregate interface view: interface route-aggregation <i>interface-number</i> 	<p>Use either command.</p> <p>Perform this step also on the intermediate device.</p> <p>The 5500 SI switch series does not support Layer 3 aggregate interfaces.</p>
4. Configure the aggregation group to work in dynamic aggregation mode.	link-aggregation mode dynamic	<p>By default, an aggregation group operates in static aggregation mode.</p> <p>Perform this step also on the intermediate device.</p>
5. Enable LACP MAD.	mad enable	By default, LACP MAD is disabled.
6. Return to system view.	quit	N/A
7. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
8. Assign the Ethernet interface to the specified aggregation group.	port link-aggregation group <i>number</i>	Perform this step also on the intermediate device.

Configuring BFD MAD

When you configure BFD MAD, follow these guidelines:

- You cannot enable BFD MAD detection on VLAN-interface 1.
- Do not use the BFD MAD VLAN for any other purpose. No Layer 2 or Layer 3 features, including ARP and LACP, can work on the BFD MAD-enabled VLAN interface or any port in the VLAN. If you configure any other feature on the VLAN, neither the configured feature nor the BFD MAD function can work correctly.
- Do not enable the spanning tree function on any port in a BFD MAD VLAN. The MAD function is mutually exclusive with the spanning tree function.
- Do not bind a BFD MAD-enabled VLAN interface to any VPN. The MAD function is mutually exclusive with VPN.
- To avoid anomalies, do not assign the BFD MAD-enabled VLAN interface a common IP address, for example, an IP address configured with the **ip address** command or a VRRP virtual IP address.
- If an intermediate device is used, you must assign the ports of the BFD MAD links to the BFD MAD VLAN on the device.

To configure BFD MAD:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a VLAN dedicated to BFD MAD.	vlan <i>vlan-id</i>	The default VLAN on the switch is VLAN 1.
3. Return to system view.	quit	N/A
4. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
5. Assign the port to the BFD MAD VLAN.	<ul style="list-style-type: none"> Assign the port to the VLAN as an access port: port access vlan <i>vlan-id</i> Assign the port to the VLAN as a trunk port: port trunk permit vlan <i>vlan-id</i> Assign the port to the VLAN as a hybrid port: port hybrid vlan <i>vlan-id</i> { tagged untagged } 	<p>Choose one command depending on the port type.</p> <p>BFD MAD detection has no requirement for the link type of the detection port.</p> <p>The default link type of a port is access.</p>
6. Return to system view.	quit	N/A
7. Enter VLAN interface view.	interface vlan-interface <i>interface-number</i>	N/A
8. Enable BFD MAD.	mad bfd enable	By default, BFD MAD is disabled.
9. Configure a MAD IP address for the specified member on the VLAN interface.	mad ip address <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } member <i>member-id</i>	<p>By default, no MAD IP address is configured on any VLAN interface.</p> <p>The MAD IP address must not on the same subnet as any IP address configured on any member device.</p>

NOTE:

After a BFD MAD-enabled IRF fabric splits, route conflict messages (for example, %May 5 16:15:47:733 2010 ARP/3/ROUTECONFLICT: Slot=5;Route conflict found, IP:192.168.2.1, VrfIndex:0) might appear on the fabric part that does not have the original master, because this fabric part still keeps the forwarding entries with the original master as the destination. This message does not affect forwarding. The system stops generating it after the forwarding entries are aged out.

Configuring ARP MAD

When you configure ARP MAD, follow these guidelines:

- If an intermediate device is used, you can use common data links as ARP MAD links. If no intermediate device is used, set up dedicated ARP MAD links between IRF member devices.
- Use a VLAN dedicated to ARP MAD.
- If an intermediate device is used, do the following:
 - Run the spanning tree feature between the IRF fabric and the intermediate device.

- Enable the IRF fabric to change its bridge MAC address as soon as the master leaves.
- Create the ARP MAD VLAN and assign the ports on the ARP MAD links to the VLAN.
- If the intermediate device is in an IRF fabric, assign this fabric a different domain ID than the ARP MAD-enabled fabric to avoid false detection of IRF partition.

To configure ARP MAD:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Assign a domain ID to the IRF fabric.	irf domain <i>domain-id</i>	The default IRF domain ID is 0.
3. Create a VLAN dedicated to ARP MAD.	vlan <i>vlan-id</i>	The default VLAN on the device is VLAN 1.
4. Return to system view.	quit	N/A
5. Enter Layer 2 Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
6. Assign the port to the ARP MAD VLAN.	<ul style="list-style-type: none"> • Assign the port to the VLAN as an access port: port access vlan <i>vlan-id</i> • Assign the port to the VLAN as a trunk port: port trunk permit vlan <i>vlan-id</i> • Assign the port to the VLAN as a hybrid port: port hybrid vlan <i>vlan-id</i> { tagged untagged } 	<p>Choose one command depending on the port type.</p> <p>ARP MAD detection has no requirement for the link type.</p> <p>The default link type of a port is access.</p>
7. Return to system view.	quit	N/A
8. Enter VLAN interface view.	interface <i>vlan-interface</i> <i>vlan-interface-id</i>	N/A
9. Assign the interface an IP address.	ip address <i>ip-address</i> { <i>mask</i> <i>mask-length</i> }	By default, no IP address is assigned to any VLAN interface.
10. Enable ARP MAD.	mad arp enable	By default, ARP MAD is disabled.

Excluding a port from the shutdown action upon detection of multi-active collision

When the IRF fabric transits to the Recovery state, all ports but the console and physical IRF ports by default automatically shut down.

You can exclude a port from the shutdown action for management or other special purposes. For example:

- Exclude a port from the shutdown action, so you can telnet to the port for managing the switch.
- Exclude a VLAN interface and its Layer 2 ports from the shutdown action, so you can log in through the VLAN interface.

CAUTION:

Excluding a VLAN interface and its Layer 2 ports from the shutdown action introduces IP collision risks because the VLAN interface might be active on both the IRF fabric in Active state and the IRF fabric in Recovery state.

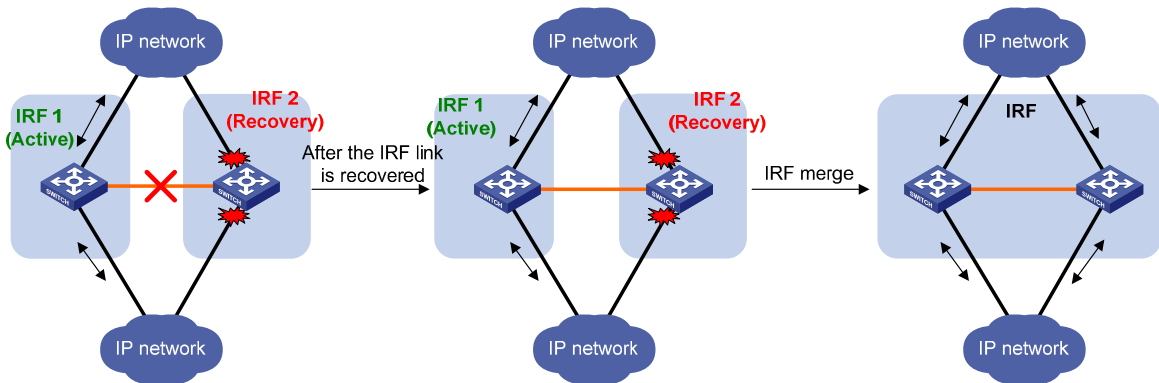
To configure a port to not shut down when the IRF fabric transits to the recovery state:

Step	Command	Remarks
1. Enter system view.	<code>system-view</code>	N/A
2. Configure a port to not shut down when the IRF fabric transits to the Recovery state.	<code>mad exclude interface interface-type interface-number</code>	By default, when an IRF fabric transits to the Recovery state, all its network ports except the IRF physical ports and console port are shut down.

Recovering an IRF fabric

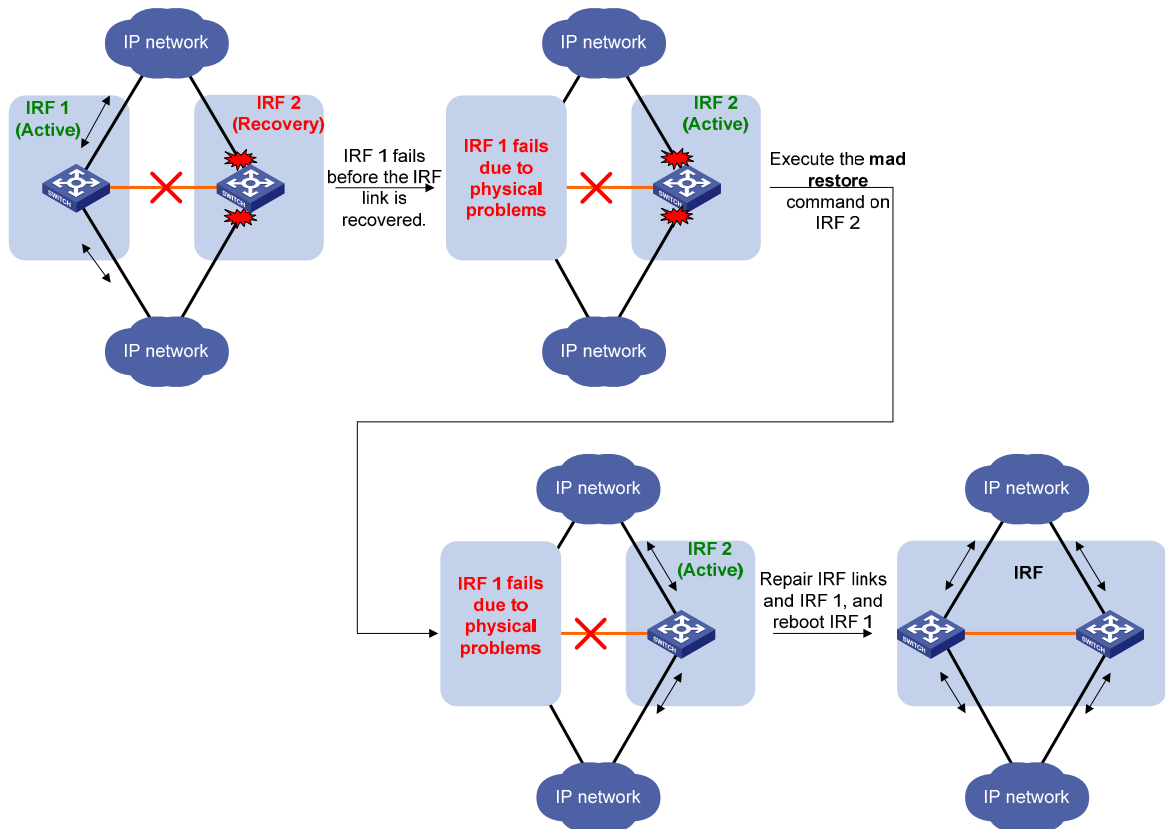
After the failed IRF link between two split IRF fabrics is recovered, log in to the Recovery-state IRF fabric, and use the **reboot** command to reboot all its members. After these member switches join the Active-state IRF fabric as subordinates, IRF merge is complete, as shown in Figure 12.

Figure 12 Recovering the IRF fabric



If the Active-state fabric has failed, for example, because of device or link failures, before the IRF link is recovered (see Figure 13), use the **mad restore** command on the Recovery-state fabric to change its state to Active for forwarding traffic. After you repair the IRF link, the two parts merge into a unified IRF fabric.

Figure 13 Active-state IRF fabric fails before the IRF link is recovered



To manually recover an IRF fabric in Recovery state:

Step	Command
1. Enter system view.	system-view
2. Change the state of the IRF fabric from Recovery to Active.	mad restore

After the IRF fabric is recovered, all ports that have been shut down by MAD automatically come up.

Displaying and maintaining an IRF fabric

Task	Command	Remarks
Display information about all IRF members.	display irf [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the IRF fabric topology.	display irf topology [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the IRF settings that take effect at the next startup.	display irf configuration [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display the load sharing mode for IRF links.	<code>display irf-port load-sharing mode</code> [<code>irf-port</code> [<code>member-id/port-number</code>]] [{ <code>begin</code> <code>exclude</code> <code>include</code> } <code>regular-expression</code>]	Available in any view
Display the master/subordinate switchover state of IRF members.	<code>display switchover state</code> [<code>slot</code> <code>member-id</code>] [{ <code>begin</code> <code>exclude</code> <code>include</code> } <code>regular-expression</code>]	Available in any view
Display MAD configuration.	<code>display mad</code> [<code>verbose</code>] [{ <code>begin</code> <code>exclude</code> <code>include</code> } <code>regular-expression</code>]	Available in any view

Configuration examples

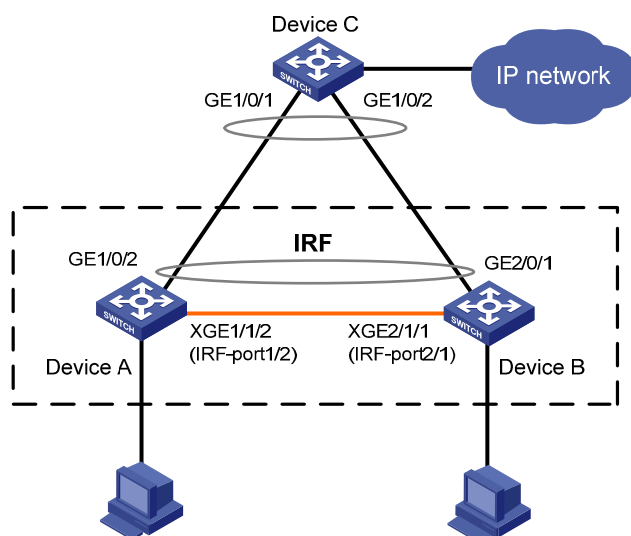
This section provides IRF configuration examples for IRF fabrics that use different MAD mechanisms.

LACP MAD-enabled IRF configuration example

Network requirements

Set up a two-member IRF fabric at the access layer of the enterprise network in Figure 14. Configure LACP MAD in the IRF fabric, because the IRF fabric has a multi-chassis aggregate link to Device C, an HP device that supports extended LACP.

Figure 14 Network diagram



Configuration procedure

This example assumes that the system names of Device A, Device B and Device C are **DeviceA**, **DeviceB**, and **DeviceC** respectively before the IRF fabric is formed.

1. Assign member IDs:
Keep the default member ID of Device A unchanged.

Change the member ID of Device B to 2.

```
<DeviceB> system-view
[DeviceB] irf member 1 renumber 2
Warning: Renumbering the switch number may result in configuration change or loss.
Continue? [Y/N]:y
[DeviceB]
```

2. Power off the devices, connect IRF links as shown in [Figure 14](#), and power on the two devices.
3. Configure IRF port bindings:

Create IRF port 2 on Device A, bind Ten-GigabitEthernet 1/1/2 to the IRF port, and save the configuration.

```
<DeviceA> system-view
[DeviceA] interface ten-gigabitethernet 1/1/2
[DeviceA-Ten-GigabitEthernet1/1/2] shutdown
[DeviceA] irf-port 1/2
[DeviceA-irf-port1/2] port group interface ten-gigabitethernet 1/1/2
[DeviceA-irf-port1/2] quit
[DeviceA] interface ten-gigabitethernet 1/1/2
[DeviceA-Ten-GigabitEthernet1/1/2] undo shutdown
[DeviceA-Ten-GigabitEthernet1/1/2] save
```

Create IRF port 1 on Device B, bind Ten-GigabitEthernet 2/1/1 to the IRF port, and save the configuration.

```
<DeviceB> system-view
[DeviceB] interface ten-gigabitethernet 2/1/1
[DeviceB-Ten-GigabitEthernet2/1/1] shutdown
[DeviceB] irf-port 2/1
[DeviceB-irf-port2/1] port group interface ten-gigabitethernet 2/1/1
[DeviceB-irf-port2/1] quit
[DeviceB] interface ten-gigabitethernet 2/1/1
[DeviceB-Ten-GigabitEthernet2/1/1] undo shutdown
[DeviceB-Ten-GigabitEthernet2/1/1] save
```

Activate IRF port configuration on Device A.

```
[DeviceA-Ten-GigabitEthernet1/1/2] quit
[DeviceA] irf-port-configuration active
```

Activate IRF port configuration on Device B.

```
[DeviceB-Ten-GigabitEthernet2/1/1] quit
[DeviceB] irf-port-configuration active
```

After the IRF port configuration is activated, the two devices automatically elect a master. In this example, Device A is the master. Device B automatically reboots and joins the Device A as a subordinate switch to form an IRF fabric. The system name of the IRF fabric is **DevicieA**.

4. Configure LACP MAD:

Create a dynamic aggregation interface and enable LACP MAD. Because LACP MAD is not run between IRF domains, you can use the default value 0.

```
<DeviceA> system-view
[DeviceA] interface bridge-aggregation 2
[DeviceA-Bridge-Aggregation2] link-aggregation mode dynamic
[DeviceA-Bridge-Aggregation2] mad enable
You need to assign a domain ID (range: 0-4294967295)
```

```

[Current domain is: 0]:
The assigned domain ID is: 0
Info: MAD LACP only enable on dynamic aggregation interface.
# Add GigabitEthernet 1/0/1 and GigabitEthernet 2/0/1 to the aggregation interface.
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] port link-aggregation group 2
[DeviceA-GigabitEthernet1/0/1] quit
[DeviceA] interface gigabitethernet 2/0/1
[DeviceA-GigabitEthernet2/0/1] port link-aggregation group 2

```

5. Configure Device C as the intermediate device:

```

# Create a dynamic aggregation interface.
<DeviceC> system-view
[DeviceC] interface bridge-aggregation 2
[DeviceC-Bridge-Aggregation2] link-aggregation mode dynamic
[DeviceC-Bridge-Aggregation2] quit
# Add GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 to the aggregation interface.
[DeviceC] interface gigabitethernet 1/0/1
[DeviceC-GigabitEthernet1/0/1] port link-aggregation group 2
[DeviceC-GigabitEthernet1/0/1] quit
[DeviceC] interface gigabitethernet 1/0/2
[DeviceC-GigabitEthernet1/0/2] port link-aggregation group 2

```

BFD MAD-enabled IRF configuration example

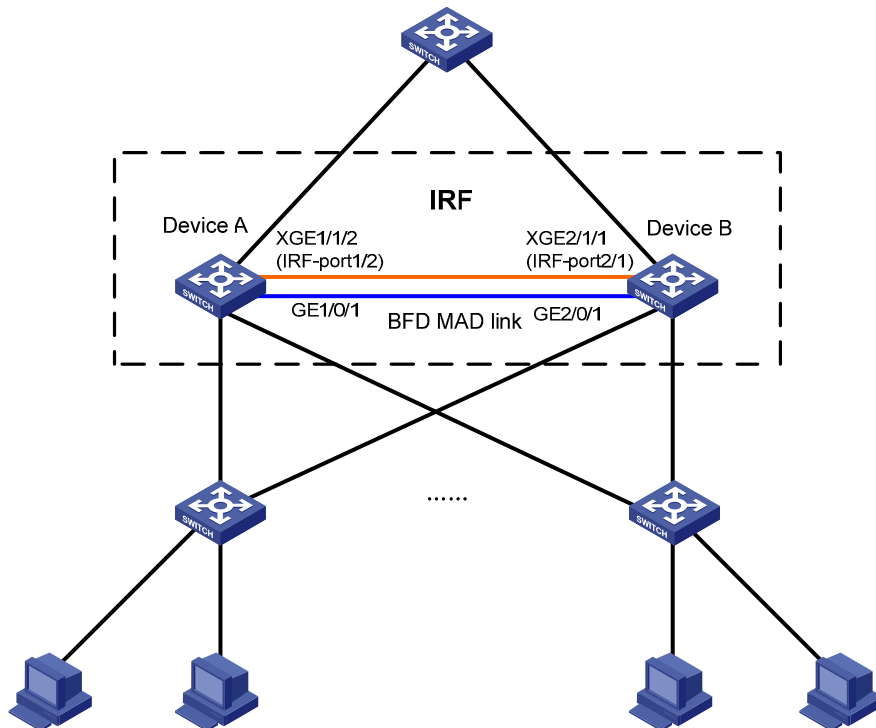
This configuration example applies only to the 5500 EI switches.

Network requirements

Set up an IRF fabric at the distribution layer of the network in [Figure 15](#). Configure BFD MAD in the IRF fabric and set up BFD MAD links between the member devices. Disable the spanning tree function on the ports used for BFD MAD, because the two features conflict with each other.

Assign the highest member priority to Device A so it can be elected as the master.

Figure 15 Network diagram



Configuration procedure

This example assumes that the system names of Device A and Device B are **DeviceA** and **DeviceB** respectively before the IRF fabric is formed.

1. Assign member IDs:

Keep the default member ID of Device A unchanged.

Change the member ID of Device B to 2.

```
<DeviceB> system-view
[DeviceB] irf member 1 renumber 2
Warning: Renumbering the switch number may result in configuration change or loss.
Continue? [Y/N]:y
[DeviceB]
```

2. Power off the member devices, connect IRF links as shown in Figure 15, and power on the two devices.

3. Configure IRF port bindings:

Create IRF port 2 on Device A, bind Ten-GigabitEthernet 1/1/2 to the IRF port, and save the configuration.

```
<DeviceA> system-view
[DeviceA] interface ten-gigabitethernet 1/1/2
[DeviceA-Ten-GigabitEthernet1/1/2] shutdown
[DeviceA] irf-port 1/2
[DeviceA-irf-port1/2] port group interface ten-gigabitethernet 1/1/2
[DeviceA-irf-port1/2] quit
[DeviceA] interface ten-gigabitethernet 1/1/2
[DeviceA-Ten-GigabitEthernet1/1/2] undo shutdown
[DeviceA-Ten-GigabitEthernet1/1/2] save
```

Create IRF port 1 on Device B, bind Ten-GigabitEthernet 2/1/1 to the IRF port, and save the configuration.

```
<DeviceB> system-view
[DeviceB] interface ten-gigabitethernet 2/1/1
[DeviceB-Ten-GigabitEthernet2/1/1] shutdown
[DeviceB] irf-port 2/1
[DeviceB-irf-port2/1] port group interface ten-gigabitethernet 2/1/1
[DeviceB-irf-port2/1] quit
[DeviceB] interface ten-gigabitethernet 2/1/1
[DeviceB-Ten-GigabitEthernet2/1/1] undo shutdown
[DeviceB-Ten-GigabitEthernet2/1/1] save
```

Activate IRF port configuration on Device A.

```
[DeviceA-Ten-GigabitEthernet1/1/2] quit
[DeviceA] irf-port-configuration active
```

Activate IRF port configuration on Device B.

```
[DeviceB-Ten-GigabitEthernet2/1/1] quit
[DeviceB] irf-port-configuration active
```

After the IRF port configuration is activated, the two devices automatically elect a master. In this example, Device A is the master. Device B automatically reboots and joins the Device A as a subordinate switch to form an IRF fabric. The system name of the IRF fabric is **DevicieA**.

4. Configure BFD MAD:

Create VLAN 3, and add port GigabitEthernet 1/0/1 on Device A (member ID 1) and port GigabitEthernet 2/0/1 on Device B (member ID 2) to VLAN 3.

```
<DeviceA> system-view
[DeviceA] vlan 3
[DeviceA-vlan3] port gigabitethernet 1/0/1 gigabitethernet 2/0/1
[DeviceA-vlan3] quit
```

Create VLAN-interface 3 and configure a MAD IP address for each member device on the interface.

```
[DeviceA] interface vlan-interface 3
[DeviceA-Vlan-interface3] mad bfd enable
[DeviceA-Vlan-interface3] mad ip address 192.168.2.1 24 member 1
[DeviceA-Vlan-interface3] mad ip address 192.168.2.2 24 member 2
[DeviceA-Vlan-interface3] quit
```

Connect the BFD MAD links as shown in [Figure 15](#).

Disable the spanning tree function on GigabitEthernet 1/0/1 and GigabitEthernet 2/0/1.

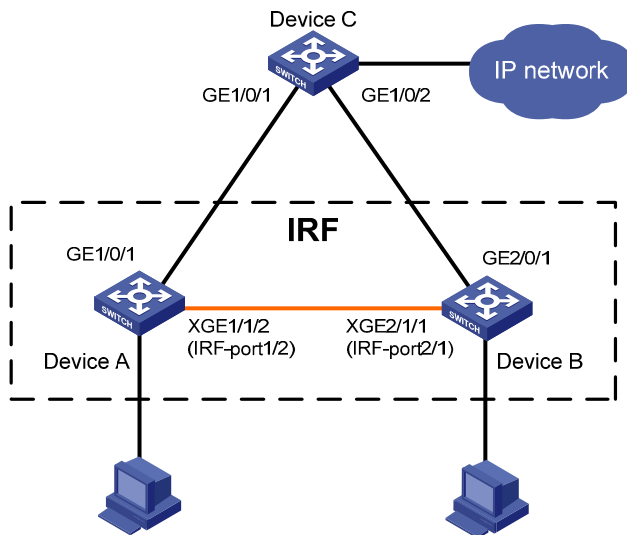
```
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-Gigabitethernet1/0/1] undo stp enable
[DeviceA-Gigabitethernet1/0/1] quit
[DeviceA] interface gigabitethernet 2/0/1
[DeviceA-Gigabitethernet2/0/1] undo stp enable
```

ARP MAD-enabled IRF configuration example

Network requirements

Set up an IRF fabric in the enterprise network in Figure 16. Configure ARP MAD in the IRF fabric and use the links connected to Device C for transmitting ARP MAD packets. To prevent loops, run the spanning tree function between Device C and the IRF fabric.

Figure 16 Network diagram



Configuration procedure

This example assumes that the system names of Device A, Device B and Device C are **DeviceA**, **DeviceB**, and **DeviceC**, respectively, before the IRF fabric is formed.

1. Assign member IDs:

Keep the default member ID of Device A unchanged.

Change the member ID of Device B to 2.

```
<DeviceB> system-view
```

```
[DeviceB] irf member 1 renumber 2
```

Warning: Renumbering the switch number may result in configuration change or loss.

Continue? [Y/N]:y

```
[DeviceB]
```

2. Power off the member devices, connect IRF links as shown in Figure 16, and power on the two devices.

3. Configure IRF port bindings:

Create IRF port 2 on Device A, and bind Ten-GigabitEthernet 1/1/2 to the IRF port, and save the configuration.

```
<DeviceA> system-view
```

```
[DeviceA] interface ten-gigabitethernet 1/1/2
```

```
[DeviceA-Ten-GigabitEthernet1/1/2] shutdown
```

```
[DeviceA] irf-port 1/2
```

```
[DeviceA-irf-port1/2] port group interface ten-gigabitethernet 1/1/2
```

```
[DeviceA-irf-port1/2] quit
```

```
[DeviceA] interface ten-gigabitethernet 1/1/2
```



```
[DeviceA-Ten-GigabitEthernet1/1/2] undo shutdown
```

```
[DeviceA-Ten-GigabitEthernet1/1/2] save
```

Create IRF port 1 on Device B, bind Ten-GigabitEthernet 2/1/1 to the IRF port, and save the configuration.

```
<DeviceB> system-view
```

```
[DeviceB] interface ten-gigabitethernet 2/1/1
```

```
[DeviceB-Ten-GigabitEthernet2/1/1] shutdown
```

```
[DeviceB] irf-port 2/1
```

```
[DeviceB-irf-port2/1] port group interface ten-gigabitethernet 2/1/1
```

```
[DeviceB-irf-port2/1] quit
```

```
[DeviceB] interface ten-gigabitethernet 2/1/1
```

```
[DeviceB-Ten-GigabitEthernet2/1/1] undo shutdown
```

```
[DeviceB-Ten-GigabitEthernet2/1/1] save
```

Activate IRF port configuration on Device A.

```
[DeviceA-Ten-GigabitEthernet1/1/2] quit
```

```
[DeviceA] irf-port-configuration active
```

Activate IRF port configuration on Device B.

```
[DeviceB-Ten-GigabitEthernet2/1/1] quit
```

```
[DeviceB] irf-port-configuration active
```

After the IRF port configuration is activated, the two devices automatically elect a master. In this example, Device A is the master. Device B automatically reboots and joins the Device A as a subordinate switch to form an IRF fabric. The system name of the IRF fabric is **DevicieA**.

4. Configure ARP MAD:

Enable MSTP globally on the IRF fabric to prevent loops.

```
<DeviceA> system-view
```

```
[DeviceA] stp enable
```

Connect the ARP MAD links as shown in [Figure 16](#).

Configure the IRF fabric to change its bridge MAC address as soon as the master leaves.

```
[DeviceA] undo irf mac-address persistent
```

Create VLAN 3, and add port GigabitEthernet 1/0/1 (on Device A) and port GigabitEthernet 2/0/1 (on Device B) to VLAN 3.

```
[DeviceA] vlan 3
```

```
[DeviceA-vlan3] port gigabitethernet 1/0/1 gigabitethernet 2/0/1
```

```
[DeviceA-vlan3] quit
```

Create VLAN-interface 3, assign it an IP address, and enable ARP MAD on the interface. Because ARP MAD is not run between IRF domains, you can use the default value 0.

```
[DeviceA] interface vlan-interface 3
```

```
[DeviceA-Vlan-interface3] ip address 192.168.2.1 24
```

```
[DeviceA-Vlan-interface3] mad arp enable
```

```
You need to assign a domain ID (range: 0-4294967295)
```

```
[Current domain is: 0]:
```

```
The assigned domain ID is: 0
```

5. Configure Device C:

Enable MSTP globally on Device C to prevent loops.

```
<DeviceC> system-view
```

```
[DeviceC] stp enable
```

```
# Create VLAN 3, and add GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 to VLAN 3.  
[DeviceC] vlan 3  
[DeviceC-vlan3] port gigabitethernet 1/0/1 gigabitethernet 1/0/2  
[DeviceC-vlan3] quit
```

Index

[A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [M](#) [P](#) [S](#)

A

- Accessing the IRF fabric, [17](#)
- Application scenario, [1](#)
- Assigning a member ID to each IRF member switch, [14](#)
- Assigning an IRF domain ID to the IRF fabric, [18](#)

B

- Basic concepts, [2](#)
- Binding physical ports to IRF ports, [16](#)

C

- Configuration examples, [29](#)
- Configuration synchronization mechanism, [6](#)
- Configuring a member switch description, [19](#)
- Configuring IRF bridge MAC persistence, [20](#)
- Configuring IRF link load sharing mode, [19](#)
- Configuring MAD, [22](#)
- Connecting physical IRF ports, [15](#)

D

- Displaying and maintaining an IRF fabric, [28](#)

E

- Enabling software auto-update for system software image synchronization, [21](#)

F

- File system naming conventions, [5](#)

G

- General restrictions and configuration guidelines, [11](#)

H

- Hardware compatibility, [1](#)

I

- Interface naming conventions, [4](#)
- IRF benefits, [1](#)
- IRF multi-active detection, [6](#)

M

- Master election, [6](#)

P

- Planning the IRF fabric setup, [13](#)

S

- Setting the IRF link down report delay, [22](#)
- Setup and configuration task list, [12](#)
- Specifying a priority for each member switch, [14](#)

Contents

Configuring Ethernet interfaces	1
Ethernet interface naming conventions	1
Performing general configurations	1
Configuring a combo interface.....	1
Configuring basic settings of an Ethernet interface.....	2
Shutting down an Ethernet interface	3
Configuring flow control on an Ethernet interface.....	3
Configuring link state change suppression on an Ethernet interface.....	4
Configuring loopback testing on an Ethernet interface.....	5
Configuring the link mode of an Ethernet interface (available only on the 5500 EI).....	6
Configuring jumbo frame support	6
Enabling the auto power-down function on an Ethernet interface	7
Configuring a Layer 2 Ethernet interface	7
Layer 2 Ethernet interface configuration task list	7
Configuring a port group	8
Configuring a MAC Address for an Ethernet port (available only on the 5500 SI).....	9
Setting speed options for auto negotiation on an Ethernet interface.....	9
Configuring storm suppression	10
Setting the statistics polling interval.....	11
Enabling loopback detection on an Ethernet interface	11
Setting the MDI mode of an Ethernet interface	13
Enabling bridging on an Ethernet interface	14
Testing the cable connection of an Ethernet interface.....	14
Configuring storm control on an Ethernet interface	15
Setting the MTU for a Layer 3 Ethernet interface (available only on the 5500 EI).....	16
Displaying and maintaining an Ethernet interface	17
Configuring loopback and null interfaces	18
Configuring a loopback interface	18
Introduction to the loopback interface.....	18
Configuration procedure	18
Configuring a null interface	19
Introduction to the null interface.....	19
Configuration procedure	19
Displaying and maintaining loopback and null interfaces.....	19
Bulk configuring interfaces	21
Configuration guidelines	21
Configuration procedure.....	21
Configuring the MAC address table	22
Overview.....	22
How a MAC address table entry is created.....	22
Types of MAC address table entries	23
MAC address table-based frame forwarding	23
Configuring static, dynamic, and blackhole MAC address table entries	23
Adding or modifying a static, dynamic, or blackhole MAC address table entry in system view	23
Adding or modifying a static or dynamic MAC address table entry in interface view	24
Disabling MAC address learning.....	24
Disabling global MAC address learning.....	24

Disabling MAC address learning on ports	24
Disabling MAC address learning on a VLAN (available only on the 5500 EI)	25
Configuring the aging timer for dynamic MAC address entries	25
Configuring the MAC learning limit on ports	26
Enabling MAC address roaming	26
Displaying and maintaining MAC address tables	28
MAC address table configuration example	28
Network requirements	28
Configuration procedure	29
Configuring MAC Information	30
Overview	30
Introduction to MAC Information	30
How MAC Information works	30
Configuring MAC Information	30
Enabling MAC Information globally	30
Enabling MAC Information on an interface	30
Configuring MAC Information mode	31
Configuring the interval for sending Syslog or trap messages	31
Configuring the MAC Information queue length	31
MAC Information configuration example	32
Configuring Ethernet link aggregation	33
Overview	33
Basic concepts	33
Aggregating links in static mode	36
Aggregating links in dynamic mode	37
Load-sharing criteria for link aggregation groups	38
Configuration restrictions and guidelines	39
Ethernet link aggregation configuration task list	39
Configuring an aggregation group	39
Configuration guidelines	40
Configuring a static aggregation group	40
Configuring a dynamic aggregation group	41
Configuring an aggregate interface	43
Configuring the description of an aggregate interface	43
Configuring the MTU of a Layer 3 aggregate interface (available only on the 5500 EI)	44
Enabling link state traps for an aggregate interface	44
Limiting the number of Selected ports for an aggregation group	45
Shutting down an aggregate interface	46
Restoring the default settings for an aggregate interface	47
Configuring load sharing for link aggregation groups	47
Configuring load-sharing criteria for link aggregation groups	47
Enabling local-first load sharing for link aggregation	48
Enabling link-aggregation traffic redirection	49
Displaying and maintaining Ethernet link aggregation	50
Ethernet link aggregation configuration examples	50
Layer 2 static aggregation configuration example	51
Layer 2 dynamic aggregation configuration example	53
Layer 3 static aggregation configuration example	55
Layer 3 dynamic aggregation configuration example	56
Configuring port isolation	58
Assigning a port to the isolation group	58
Displaying and maintaining the isolation group	58
Port isolation configuration example	59

Configuring spanning tree protocols.....	60
STP.....	60
STP protocol packets.....	60
Basic concepts in STP.....	61
Calculation process of the STP algorithm.....	62
RSTP.....	67
PVST.....	67
MSTP.....	68
STP, RSTP, and PVST limitations.....	68
MSTP features.....	68
MSTP basic concepts.....	68
How MSTP works.....	72
Implementation of MSTP on devices.....	72
Protocols and standards.....	73
Spanning tree configuration task list.....	73
Configuration restrictions and guidelines.....	73
STP configuration task list.....	73
RSTP configuration task list.....	74
PVST configuration task list.....	75
MSTP configuration task list.....	76
Setting the spanning tree mode.....	77
Configuring an MST region.....	78
Configuration restrictions and guidelines.....	78
Configuration procedure.....	78
Configuring the root bridge or a secondary root bridge.....	79
Configuration restrictions and guidelines.....	79
Configuring the current device as the root bridge of a specific spanning tree.....	80
Configuring the current device as a secondary root bridge of a specific spanning tree.....	80
Configuring the device priority.....	80
Configuring the maximum hops of an MST region.....	81
Configuring the network diameter of a switched network.....	81
Configuring spanning tree timers.....	82
Configuration restrictions and guidelines.....	82
Configuration procedure.....	83
Configuring the timeout factor.....	83
Configuring the maximum port rate.....	84
Configuring edge ports.....	84
Configuration restrictions and guidelines.....	84
Configuration procedure.....	84
Configuring path costs of ports.....	85
Specifying a standard for the device to use when it calculates the default path cost.....	85
Configuring path costs of ports.....	87
Configuration example.....	87
Configuring the port priority.....	88
Configuring the port link type.....	88
Configuration restrictions and guidelines.....	88
Configuration procedure.....	88
Configuring the mode a port uses to recognize/send MSTP packets.....	89
Enabling outputting port state transition information.....	90
Enabling the spanning tree feature.....	90
Configuration restrictions and guidelines.....	90
Enabling the spanning tree feature (in STP/RSTP/MSTP mode).....	90
Enabling the spanning tree feature (in PVST mode).....	91
Performing mCheck.....	91

Performing mCheck globally	92
Performing mCheck in interface view	92
Configuring Digest Snooping	92
Configuration restrictions and guidelines	92
Configuration procedure	93
Digest Snooping configuration example	93
Configuring No Agreement Check	94
Configuration prerequisites	95
Configuration procedure	95
No Agreement Check configuration example	96
Configuring TC snooping	96
Configuration restrictions and guidelines	97
Configuration procedure	97
Configuring protection functions	98
Configuration prerequisites	98
Enabling BPDU guard	98
Enabling root guard	98
Enabling loop guard	99
Enabling TC-BPDU guard	100
Enabling BPDU drop	100
Displaying and maintaining the spanning tree	101
Spanning tree configuration examples	102
MSTP configuration example	102
PVST configuration example	105
Configuring BPDU tunneling	109
Overview	109
Background	109
BPDU tunneling implementation	110
Enabling BPDU tunneling	111
Configuration prerequisites	111
Configuration restrictions and guidelines	112
Enabling BPDU tunneling	112
Configuring destination multicast MAC address for BPDUs	112
BPDU tunneling configuration examples	113
BPDU tunneling for STP configuration example	113
BPDU tunneling for PVST configuration example	114
Configuring VLANs	116
Overview	116
VLAN fundamentals	116
VLAN types	117
Protocols and standards	118
Configuring basic VLAN settings	118
Configuration restrictions and guidelines	118
Configuration procedure	118
Configuring basic settings of a VLAN interface	119
Configuration procedure	119
VLAN interface configuration example	120
Configuring port-based VLANs	121
Introduction to port-based VLAN	121
Assigning an access port to a VLAN	122
Assigning a trunk port to a VLAN	123
Assigning a hybrid port to a VLAN	124
Port-based VLAN configuration example	125

Configuring MAC-based VLANs	127
Introduction to MAC-based VLAN	127
Configuration restrictions and guidelines	129
Configuration procedure	129
MAC-based VLAN configuration example	132
Configuring protocol-based VLANs	134
Introduction to protocol-based VLAN	134
Configuration restrictions and guidelines	134
Configuration procedure	135
Protocol-based VLAN configuration example	135
Configuring IP subnet-based VLANs	138
Configuration procedure	138
IP subnet-based VLAN configuration example	139
Displaying and maintaining VLAN	141
Configuring a super VLAN (available only on the 5500 EI)	142
Overview	142
Configuration procedure	142
Displaying and maintaining super VLAN	144
Super VLAN configuration example	144
Configuring an isolate-user-VLAN	147
Overview	147
Configuration restrictions and guidelines	148
Configuration procedure	148
Displaying and maintaining isolate-user-VLAN	149
Isolate-user-VLAN configuration example	150
Configuring a voice VLAN	153
Overview	153
OUI addresses	153
Voice VLAN assignment modes	153
Security mode and normal mode of voice VLANs	155
Configuration prerequisites	156
Configuring QoS priority settings for voice traffic on an interface	157
Configuration restrictions and guidelines	157
Configuration procedure	157
Configuring a port to operate in automatic voice VLAN assignment mode	157
Configuring a port to operate in manual voice VLAN assignment mode	158
Configuration restrictions and guidelines	158
Configuration procedure	158
Displaying and maintaining voice VLAN	159
Voice VLAN configuration examples	159
Automatic voice VLAN mode configuration example	159
Manual voice VLAN assignment mode configuration example	161
Configuring GVRP	164
Overview	164
GARP	164
GVRP	167
Protocols and standards	167
GVRP configuration task list	167
Configuring GVRP functions	168
Configuration restrictions and guidelines	168
Configuration procedure	168
Configuring the GARP timers	169

Displaying and maintaining GVRP.....	170
GVRP configuration examples	170
GVRP normal registration mode configuration example.....	170
GVRP fixed registration mode configuration example.....	172
GVRP forbidden registration mode configuration example.....	173
Configuring QinQ.....	176
Overview.....	176
Background and benefits.....	176
How QinQ works.....	176
QinQ frame structure	177
Implementations of QinQ	178
Modifying the TPID in a VLAN tag	178
Protocols and standards	179
QinQ configuration task list.....	179
Configuring basic QinQ	180
Enabling basic QinQ.....	180
Configuring VLAN transparent transmission	180
Configuring selective QinQ.....	181
Configuring an outer VLAN tagging policy	181
Configuring an inner-outer VLAN 802.1p priority mapping.....	183
Configuring inner VLAN ID substitution (available only on the 5500 EI)	184
Configuring the TPID value in VLAN tags.....	185
Configuring the TPID value on the 5500 EI.....	185
Configuring the TPID value on the 5500 SI.....	185
QinQ configuration examples	186
Basic QinQ configuration example.....	186
Selective QinQ configuration example (configuring an outer VLAN tagging policy in port-based approach)	188
Selective QinQ configuration example (configuring an outer VLAN tagging policy in QoS policy-based approach).....	191
Selective QinQ configuration example (inner VLAN ID substitution).....	194
Configuring VLAN mapping (available only on the 5500 EI).....	199
Overview.....	199
Application scenario of one-to-one and many-to-one VLAN mapping	199
Application scenario of two-to-two VLAN mapping	200
Concepts and terms	201
VLAN mapping implementations	202
VLAN mapping configuration tasks	203
Configuring one-to-one VLAN mapping	204
Configuration prerequisites	204
Configuring an uplink policy.....	204
Configuring a downlink policy	204
Configuring the customer-side port.....	205
Configuring the network-side port	205
Configuring many-to-one VLAN mapping	206
Configuration prerequisites	206
Enabling DHCP snooping.....	206
Enabling ARP detection in SVLANs	207
Configuring an uplink policy.....	207
Configuring the customer-side port.....	207
Configuring the network-side port	208
Configuring two-to-two VLAN mapping.....	208
Configuring an uplink policy for the customer-side port	209

Configuring an uplink policy for the network-side port.....	209
Configuring a downlink policy for the customer-side port.....	210
Configuring the customer-side port.....	211
Configuring the network-side port.....	211
VLAN mapping configuration examples.....	212
One-to-one and many-to-one VLAN mapping configuration example.....	212
Two-to-two VLAN mapping configuration example.....	218
Configuring LLDP.....	221
Overview.....	221
Background.....	221
Basic concepts.....	221
How LLDP works.....	225
Protocols and standards.....	226
LLDP configuration task list.....	226
Performing basic LLDP configuration.....	227
Enabling LLDP.....	227
Setting the LLDP operating mode.....	227
Setting the LLDP re-initialization delay.....	227
Enabling LLDP polling.....	228
Configuring the advertisable TLVs.....	228
Configuring the management address and its encoding format.....	229
Setting other LLDP parameters.....	230
Setting an encapsulation format for LLDPDUs.....	230
Configuring CDP compatibility.....	231
Configuration prerequisites.....	231
Configuring CDP compatibility.....	231
Enabling LLDP to automatically discover IP phones.....	232
Configuration prerequisites.....	232
Configuration procedure.....	233
Configuring LLDP to advertise a specific voice VLAN.....	233
Configuration guidelines.....	233
Configuration procedure.....	234
Dynamically advertising server-assigned VLANs through LLDP.....	235
Configuring LLDP trapping.....	235
Displaying and maintaining LLDP.....	235
LLDP configuration examples.....	236
Basic LLDP configuration example.....	236
CDP-compatible LLDP configuration example.....	239
Configuring a service loopback group (available only on the 5500 EI).....	241
Overview.....	241
Service types of service loopback groups.....	241
Requirements on service loopback ports.....	241
States of service loopback ports.....	241
Configuration restrictions and guidelines.....	242
Configuration procedure.....	242
Displaying and maintaining service loopback groups.....	243
Service loopback group configuration example.....	243
Network requirements.....	243
Configuration procedure.....	243
Configuring MVRP.....	245
Overview.....	245
Introduction to MRP.....	245
MVRP registration modes.....	247

Protocols and standards	248
MVRP configuration task list	248
Configuration prerequisites	248
Enabling MVRP	248
Configuration restrictions and guidelines	248
Configuration procedure	249
Configuring the MVRP registration mode	249
Configuring MRP timers	250
Enabling GVRP compatibility	251
Configuration restrictions and guidelines	251
Configuration procedure	251
Displaying and maintaining MVRP	251
Configuration example for MVRP in normal registration mode	252
Network requirements	252
Configuration procedure	253
Index	261

Configuring Ethernet interfaces

Ethernet interface naming conventions

The GE and 10-GE interfaces on the switches are named in the format of *interface-type A/B/C*, where the following definitions apply:

- **A**—Represents the ID of the switch in an IRF fabric. If the switch is not assigned to any IRF fabric, A uses 1.
- **B**—Represents a slot number on the switch. It uses 0 for fixed interfaces, 1 for interfaces on interface expansion card 1, and 2 for interfaces on interface expansion card 2.
- **C**—Represents the number of an interface on a slot.

NOTE:

The GE ports of the GE interface expansion card are displayed as 10-GE ports on the switch where the expansion card is installed.

For more information about the expansion cards, see *HP 5500 EI Switch Series Installation Guide* and *HP 5500 SI Switch Series Installation Guide*.

Performing general configurations

This section describes the attributes and configurations common to Layer 2 and Layer 3 Ethernet interfaces.

- For more information about the attributes and configuration of Layer 2 Ethernet interfaces, see "[Configuring a Layer 2 Ethernet interface](#)."
- For more information about the attributes and configuration of Layer 3 Ethernet interfaces, see "[Setting the MTU for a Layer 3 Ethernet interface](#)."

Configuring a combo interface

A combo interface is a logical interface that comprises one optical (fiber) port and one electrical (copper) port. The two ports share one forwarding interface, so they cannot work simultaneously. When you enable one port, the other is automatically disabled.

The fiber combo port and copper combo port are Layer 2 Ethernet interfaces. They have their own separate interface views, in which you can activate the fiber or copper combo port and configure other port attributes such as the interface rate and duplex mode.

Configuration prerequisites

Before you configure a combo interface, complete the following tasks:

- Use the **display port combo** command to identify the combo interfaces on your switch and identify the two physical ports that compose each combo interface.
- Use the **display interface** command to determine, of the two physical ports that compose a combo interface, which is the fiber combo port and which is the copper combo port. If the current port is the copper port, the output will include "Media type is twisted pair, Port hardware type is

1000_BASE_T". If the current port is the fiber port, the output will not include the information mentioned above.

Changing the active port of a combo interface

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Activate the current interface.	undo shutdown	Optional. By default, of the two ports that compose a combo interface, the one with a smaller port ID is active.

Configuring basic settings of an Ethernet interface

You can set an Ethernet interface to operate in one of the following duplex modes:

- **Full-duplex mode (full)**—Interfaces that operate in this mode can send and receive packets simultaneously.
- **Half-duplex mode (half)**—Interfaces that operate in this mode cannot send and receive packets simultaneously.
- **Auto-negotiation mode (auto)**—Interfaces that operate in this mode negotiate a duplex mode with their peers.

You can set the speed of an Ethernet interface or enable it to automatically negotiate a speed with its peer. For a 100-Mbps or 1000-Mbps Layer 2 Ethernet interface, you can also set speed options for auto negotiation. The two ends can select a speed only from the available options. For more information, see "[Setting speed options for auto negotiation on an Ethernet interface.](#)"

To configure an Ethernet interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Set the interface description.	description <i>text</i>	Optional. By default, the description of an interface is in the format of <i>interface-name</i> Interface . For example, GigabitEthernet1/0/1 Interface .

Step	Command	Remarks
4. Set the duplex mode of the interface.	duplex { auto full half }	Optional. By default, the duplex mode is auto for Ethernet interfaces. The half keyword is not applicable to Ethernet copper ports that are configured with a 1000-Mbps port speed and fiber ports.
5. Set the port speed.	speed { 10 100 1000 auto }	Optional. By default, an Ethernet interface automatically negotiates a speed with the peer. A Gigabit fiber port does not support the 10 or 100 keyword, A 10-Gigabit fiber port does not support this command
6. Restore the default settings for the interface.	default	Optional.

Shutting down an Ethernet interface

△ CAUTION:

Use this feature with caution. After you manually shut down an Ethernet interface, the Ethernet interface cannot forward packets even if it is physically connected.

You might need to shut down and then bring up an Ethernet interface to activate some configuration changes, for example, the speed or duplex mode changes.

To shut down an Ethernet interface or a group of Ethernet interfaces:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view or port group view.	<ul style="list-style-type: none"> Enter Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use any command. To shut down an Ethernet interface, enter Ethernet interface. To shut down all Ethernet interfaces in a port group, enter port group view.
3. Shut down the Ethernet interface or interfaces.	shutdown	By default, Ethernet interfaces are up.

Configuring flow control on an Ethernet interface

To avoid packet drops on a link, you can enable flow control at both ends of the link. When traffic congestion occurs at the receiving end, the receiving end sends a flow control (Pause) frame to ask the sending end to suspend sending packets

- With the **flow-control** command configured, an interface can both send and receive flow control frames: When congested, the interface sends a flow control frame to its peer. Upon receiving a flow control frame from the peer, the interface suspends sending packets.
- With the **flow-control receive enable** command configured, an interface can receive, but not send flow control frames. When the interface receives a flow control frame from its peer, it suspends sending packets to the peer. When congested, the interface cannot send flow control frames to the peer.

To handle unidirectional traffic congestion on a link, configure the **flow-control receive enable** command at one end, and the **flow-control** command at the other. To enable both ends of the link to handle traffic congestion, configure the **flow-control** command at both ends.

To enable flow control on an Ethernet interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable flow control.	<ul style="list-style-type: none"> • Enable TxRx flow control: flow-control • Enable Rx flow control: flow-control receive enable 	<p>Use either command.</p> <p>By default, Rx flow control is disabled on an Ethernet interface.</p>

Configuring link state change suppression on an Ethernet interface

An Ethernet interface has two physical link states: up and down. Each time the physical link of an interface goes up or comes down, the physical layer reports the change to the upper layers, and the upper layers handle the change, resulting in increased overhead.

To prevent physical link flapping from affecting system performance, configure link change suppression to delay the reporting of physical link state changes. When the delay expires, the interface reports any detected change.

Link change suppression does not suppress administrative up or down events. When you shut down or bring up an interface by using the **shutdown** or **undo shutdown** command, the interface reports the event to the upper layers immediately.

Link-down event suppression enables an interface to suppress link-down events and start a delay timer each time the physical link goes down. During this delay, the interface does not report the link-down event, and the **display interface brief** or **display interface** command displays the interface state as UP. If the physical link is still down when the timer expires, the interface reports the link-down event to the upper layers.

Link-up event suppression enables an interface to suppress link-up events and start a delay timer each time the physical link goes up. During this delay, the interface does not report the link-up event, and the **display interface brief** or **display interface** command displays the interface state as DOWN. If the physical link is still up when the timer expires, the interface reports the link-up event to the upper layers.

Configuring link-down event suppression

To enable an Ethernet interface to suppress link-down events:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Set a link-down event suppression interval.	link-delay <i>delay-time</i>	Link-down event suppression is disabled by default.

Configuring link-up event suppression

To configure link-up event suppression on an Ethernet interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Set a link-up event suppression interval.	link-delay <i>delay-time</i> mode up	Link-up event suppression is disabled by default.

NOTE:

The **link-delay mode up** command and the **link-delay** command supersedes each other, and whichever is configured last takes effect.

Configuring loopback testing on an Ethernet interface

If an Ethernet interface does not work normally, you can enable loopback testing on it to identify the problem. Loopback testing has the following types:

- **Internal loopback testing**—Tests all on-chip functions related to Ethernet interfaces.
- **External loopback testing**—Tests hardware of Ethernet interfaces. To perform external loopback testing on an Ethernet interface, connect a loopback plug to the Ethernet interface. The switch sends test packets out of the interface, which are expected to loop over the plug and back to the interface. If the interface fails to receive any test packet, the hardware of the interface is faulty.

An Ethernet interface in a loopback test does not forward data traffic.

Configuration restrictions and guidelines

- On an interface that is physically down, you can only perform internal loopback testing. On an interface administratively shut down, you can perform neither internal nor external loopback testing.
- During loopback testing, the Ethernet interface operates in full duplex mode. When you disable loopback testing, the interface returns to its duplex setting.
- Loopback testing is a one-time operation, and is not recorded in the configuration file.

Configuration procedure

To enable loopback testing on an Ethernet interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable loopback testing.	loopback { external internal }	Optional. Disabled by default.

Configuring the link mode of an Ethernet interface (available only on the 5500 EI)

⚠ CAUTION:

- After you change the link mode of an Ethernet interface, all the settings of the Ethernet interface are restored to their defaults under the new link mode.
- The link mode configuration for an Ethernet interface in system view and in interface view supersedes each other.

An Ethernet interface operates either in Layer 2 (bridge) or Layer 3 (route) mode. To meet networking requirements, you can use a command to set the link mode of an Ethernet interface to bridge or route.

To change the link mode of an Ethernet interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Change the link mode of Ethernet interfaces.	<ul style="list-style-type: none"> • In system view: port link-mode { bridge route } <i>interface-list</i> • In Ethernet interface view: a. interface <i>interface-type</i> <i>interface-number</i> b. port link-mode { bridge route } 	Use either approach.

Configuring jumbo frame support

An Ethernet interface might receive some frames larger than the standard Ethernet frame size (called "jumbo frames") during high-throughput data exchanges such as file transfers. Usually, an Ethernet interface discards jumbo frames. With jumbo frame support enabled, the interface can process frames larger than the standard Ethernet frame size yet within the specified range.

In interface configuration mode (Ethernet interface view or port group view), you can set the length of jumbo frames that are allowed to pass through the Ethernet interface.

- If you execute the command in Ethernet interface view, the configuration takes effect only on the interface.
- If you execute the command in port group view, the configuration takes effect on all ports in the port group.

To configure jumbo frame support in interface view or port group view:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure jumbo frame support.	jumboframe enable [<i>value</i>]	By default, the switch allows jumbo frames within 9216 bytes to pass through Ethernet interfaces. If you set the <i>value</i> argument multiple times, the latest configuration takes effect.

Enabling the auto power-down function on an Ethernet interface

To save power, enable the auto power-down function on Ethernet interfaces. An interface enters the power save mode if it has not received any packet for a certain period of time (this interval depends on the specifications of the chip, and is not configurable). When a packet arrives later, the interface enters its normal state.

To enable auto power-down on an Ethernet interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable auto power-down on an Ethernet interface.	port auto-power-down	Disabled by default.

NOTE:

When you connect an interface enabled with auto power-down to a device, if the link cannot go up properly, disable auto power-down on the interface and try again.

Configuring a Layer 2 Ethernet interface

Layer 2 Ethernet interface configuration task list

Task	Remarks
Configuring a port group	Optional Applicable to Layer 2 Ethernet interfaces
Configuring a MAC Address for an Ethernet port (available only on the 5500 S1)	Optional Applicable to Layer 2 Ethernet interfaces
Setting speed options for auto negotiation on an Ethernet interface	Optional Applicable to Layer 2 Ethernet interfaces

Task	Remarks
Configuring storm suppression	Optional Applicable to Layer 2 Ethernet interfaces
Setting the statistics polling interval	Optional Applicable to Layer 2 Ethernet interfaces
Enabling loopback detection on an Ethernet interface	Optional Applicable to Layer 2 Ethernet interfaces
Setting the MDI mode of an Ethernet interface	Optional Applicable to Layer 2 Ethernet interfaces
Enabling bridging on an Ethernet interface	Optional Applicable to Layer 2 Ethernet interfaces
Testing the cable connection of an Ethernet interface	Optional Applicable to Layer 2 Ethernet interfaces
Configuring storm control on an Ethernet interface	Optional Applicable to Layer 2 Ethernet interfaces

Configuring a port group

Some interfaces on your switch might use the same set of settings. To configure these interfaces in bulk rather than one by one, you can assign them to a port group.

You create port groups manually. All settings made for a port group apply to all the member ports of the group.

Even though the settings are made on the port group, they are saved on each interface basis rather than on a port group basis. You can only view the settings in the view of each interface by using the **display current-configuration** or **display this** command.

To configure a manual port group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a manual port group and enter manual port group view.	port-group manual <i>port-group-name</i>	N/A
3. Assign Ethernet interfaces to the manual port group.	group-member <i>interface-list</i>	N/A
4. Configure jumbo frame support.	jumboframe enable [<i>value</i>]	By default, the switch allows jumbo frames within 9216 bytes to pass through Ethernet interfaces. If you set the <i>value</i> argument multiple times, the latest configuration takes effect.
5. Enable auto power-down.	port auto-power-down	Disabled by default.

Configuring a MAC Address for an Ethernet port (available only on the 5500 SI)

You can use this function to manually configure a MAC address for an Ethernet port. After that, when the switch sends packets of Layer 2 protocols (including cluster, DLDAP, HABP, LACP, LLDP, MSTP, NDP, NTDP, and GVRP) out of this port, it can use the MAC address of this port as the source MAC address of these protocol packets.

Configuring a MAC address for an Ethernet port does not affect the normal forwarding of service packets.

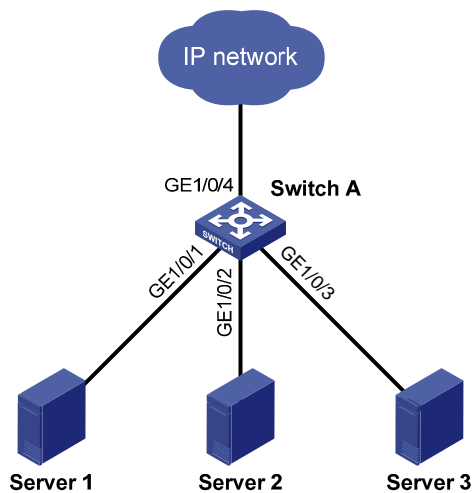
To configure a MAC Address for an Ethernet port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet port view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure a MAC address for the current port.	mac-address <i>mac-address</i>	The default MAC address of an Ethernet port varies with device models.

Setting speed options for auto negotiation on an Ethernet interface

Speed auto negotiation enables an Ethernet interface to negotiate with its peer for the highest speed that both ends support by default. You can narrow down the speed option list for negotiation.

Figure 1 Speed auto negotiation application scenario



As shown in Figure 1, all ports on Switch A are operating in speed auto negotiation mode, with the highest speed of 1000 Mbps. If the transmission rate of each server in the server cluster is 1000 Mbps, their total transmission rate will exceed the capability of port GigabitEthernet 1/0/4, the port providing access to the Internet for the servers.

To avoid congestion on GigabitEthernet 1/0/4, set 100 Mbps as the only option available for speed negotiation on port GigabitEthernet 1/0/1, GigabitEthernet 1/0/2, and GigabitEthernet 1/0/3. As a result, the transmission rate on each port connected to a server is limited to 100 Mbps.

To set speed options for auto negotiation on an Ethernet interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type interface-number</i>	N/A
3. Set speed options for auto negotiation.	speed auto { 10 100 1000 } *	Optional.

NOTE:

- This function is available only for Layer 2 copper GE ports that support speed auto negotiation.
- The **speed** and **speed auto** commands supersede each other, and whichever is configured last takes effect.

Configuring storm suppression

You can use the storm suppression function to limit the size of a particular type of traffic (broadcast, multicast, or unknown unicast traffic) as a whole globally in system view or on a per-interface basis in Ethernet interface view or port group view.

In interface or port group view, you set the maximum size of broadcast, multicast or unknown unicast traffic allowed to pass through an interface or each interface in a port group. When the broadcast, multicast, or unknown unicast traffic on the interface exceeds this threshold, the system discards packets until the traffic drops below this threshold.

The storm suppression thresholds configured for an Ethernet interface might become invalid if you enable the storm control function for the interface. For information about the storm control function, see "[Configuring storm control on an Ethernet interface.](#)"

To set storm suppression thresholds on one or multiple Ethernet interfaces:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view or port group view.	<ul style="list-style-type: none"> • Enter Ethernet interface view: interface <i>interface-type interface-number</i> • Enter port group view: port-group manual <i>port-group-name</i> 	<p>Use either command.</p> <p>To configure storm suppression on an Ethernet interface, enter Ethernet interface view.</p> <p>To configure storm suppression on a group of Ethernet interfaces, enter port group view.</p>
3. Set the broadcast suppression threshold ratio.	broadcast-suppression { <i>ratio</i> pps <i>max-pps</i> kbps <i>max-kbps</i> }	<p>Optional.</p> <p>By default, all broadcast traffic is allowed to pass through.</p>
4. Set the multicast suppression threshold ratio.	multicast-suppression { <i>ratio</i> pps <i>max-pps</i> kbps <i>max-kbps</i> }	<p>Optional.</p> <p>By default, all multicast traffic is allowed to pass through.</p>

Step	Command	Remarks
5. Set the unknown unicast suppression threshold ratio.	unicast-suppression { <i>ratio</i> pps <i>max-pps</i> kbps <i>max-kbps</i> }	Optional. By default, all unknown unicast traffic is allowed to pass through.

NOTE:

For an Ethernet interface that belongs to a port group, if you set a traffic suppression threshold for the interface in both Ethernet interface view and port group view, the threshold configured last takes effect.

Setting the statistics polling interval

To set the statistics polling interval globally or on an Ethernet interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Set the statistics polling interval on the Ethernet interface.	flow-interval <i>interval</i>	Optional. The default interface statistics polling interval is 300 seconds.

To display the interface statistics collected in the last polling interval, use the **display interface** command.

To clear interface statistics, use the **reset counters interface** command.

Enabling loopback detection on an Ethernet interface

If a switch receives a packet that it sent, a loop has occurred to the switch. Loops might cause broadcast storms, which degrade network performance. You can use this feature to detect whether a loop has occurred.

Depending on whether the receiving interface is the same as the sending interface, loops fall into the following types:

- **Single-port loopback**—Occurs when an interface receives a packet that it sent out and the receiving interface is the same as the sending interface, as shown in [Figure 2](#).
- **Multi-port loopback**—Occurs when a switch receives a packet that it sent out but the receiving interface might not be the sending interface, as shown in [Figure 3](#).

Figure 2 Single-port loopback

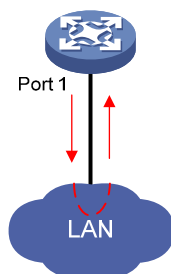
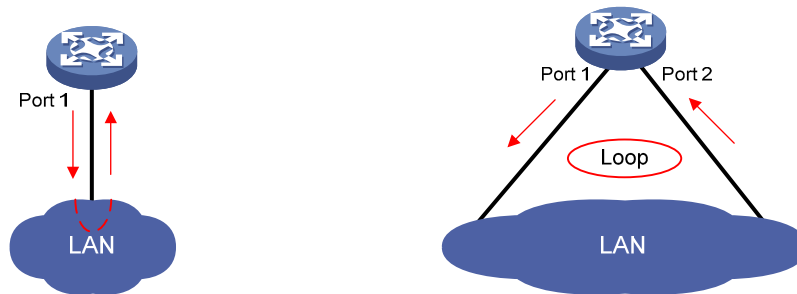


Figure 3 Multi-port loopback



You can enable loopback detection to detect loops on an interface and, if the interface supports the **loopback-detection action** command, configure the protective action to take on the receiving interface when a loop is detected, for example, to shut down the interface. Depending on whether a protective action is configured, the switch takes the actions in [Table 1](#) to alleviate the impact of the loop condition.

Table 1 Actions to take upon detection of a loop condition

Port type	Actions	
	No protective action is configured	A protective action is configured
Access port	<ul style="list-style-type: none"> Place the receiving interface in controlled mode. The interface does not receive or send packets. Generate traps and log messages. Delete all MAC address entries of the interface. 	<ul style="list-style-type: none"> Perform the configured protective action. Generate traps and log messages. Delete all MAC address entries of the interface.
Hybrid or trunk port	<ul style="list-style-type: none"> Generate traps and log messages. If loopback detection control is enabled, place the receiving interface in controlled mode. The interface does not receive or send packets. Delete all MAC address entries of the interface. 	<ul style="list-style-type: none"> Generate traps and log messages. If loopback detection control is enabled, take the configured protective action on the interface. Delete all MAC address entries of the interface.

Configuration restrictions and guidelines

- To use loopback detection on an Ethernet interface, you must enable the function both globally and on the interface.
- To disable loopback detection on all interfaces, run the **undo loopback-detection enable** command in system view.
- To enable a hybrid or trunk port to take the administratively specified protective action, you must use the **loopback-detection control enable** command on the port.
- When you change the link type of an Ethernet interface by using the **port link-type** command, the switch removes the protective action configured on the interface. For more information about the **port link-type** command, see *Layer 2—LAN Switching Command Reference*.

Configuration procedure

To configure loopback detection:

Step	Command	Remarks
1.	Enter system view. system-view	N/A

Step	Command	Remarks
2. Enable global loopback detection.	loopback-detection enable	Disabled by default.
3. Enable multi-port loopback detection.	loopback-detection multi-port-mode enable	Optional. By default, multi-port loopback detection is disabled, and the switch can only detect single-port loopback.
4. Set the loopback detection interval.	loopback-detection interval-time <i>time</i>	Optional. 30 seconds by default.
5. Enter Ethernet interface view or port group view.	<ul style="list-style-type: none"> Enter Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command. To configure loopback detection on one interface, enter Ethernet interface view. To configure loopback detection on a group of Ethernet interfaces, enter port group view.
6. Enable loopback detection on the interface.	loopback-detection enable	Disabled by default.
7. Enable loopback detection control on a trunk port or a hybrid port.	loopback-detection control enable	Optional. Disabled by default.
8. Enable loopback detection in all the VLANs on the trunk or hybrid port.	loopback-detection per-vlan enable	Optional. By default, a trunk or hybrid port performs loopback detection only in its port VLAN ID (PVID).
9. Set the protective action to take on the interface when a loop is detected.	loopback-detection action { no-learning semi-block shutdown }	Optional. By default, a looped interface does not receive or send packets; the system generates traps and log messages, and deletes all MAC address entries of the looped interface. With the shutdown keyword specified, the switch shuts down the looped ports and set their physical state to Loop down. When a looped port recovers, you must use the undo shutdown command to restore its forwarding capability.

Setting the MDI mode of an Ethernet interface

ⓘ IMPORTANT:

Fiber ports do not support the MDI mode setting.

You can use both crossover and straight-through Ethernet cables to connect copper Ethernet interfaces. To accommodate these types of cables, a copper Ethernet interface can operate in one of the following Medium Dependent Interface (MDI) modes:

- Across mode

- Normal mode
- Auto mode

A copper Ethernet interface uses an RJ-45 connector, which comprises eight pins, each of which plays a dedicated role. For example, pins 1 and 2 transmit signals, and pins 3 and 6 receive signals. The pin role varies by the MDI modes as follows:

- In normal mode, pins 1 and 2 are transmit pins, and pins 3 and 6 are receive pins.
- In across mode, pins 1 and 2 are receive pins, and pins 3 and 6 are transmit pins.
- In auto mode, the interface negotiates pin roles with its peer.

To enable the interface to communicate with its peer, make sure that its transmit pins are connected to the remote receive pins. If the interface can detect the connection cable type, set the interface in auto MDI mode. If not, set its MDI mode by using the following guidelines:

- When a straight-through cable is used, set the interface to operate in the MDI mode different than its peer.
- When a crossover cable is used, set the interface to operate in the same MDI mode as its peer, or set either end to operate in auto mode.

To set the MDI mode of an Ethernet interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Set the MDI mode of the Ethernet interface.	mdi { across auto normal }	Optional. By default, a copper Ethernet interface operates in auto mode to negotiate pin roles with its peer.

Enabling bridging on an Ethernet interface

When an incoming packet arrives, the device looks up the destination MAC address of the packet in the MAC address table. If an entry is found, but the outgoing interface is the same as the receiving interface, the device discards the packet.

To enable the device to forward such packets rather than drop them, enable the bridging function on the Ethernet interface.

To enable bridging on an Ethernet interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable bridging on the Ethernet interface.	port bridge enable	Disabled by default.

Testing the cable connection of an Ethernet interface

❗ **IMPORTANT:**

- Fiber ports do not support this feature.
 - If the link of an Ethernet port is up, testing its cable connection will cause the link to come down and then go up.
-

You can test the cable connection of an Ethernet interface for a short or open circuit. The switch displays cable test results within five seconds. If any fault is detected, the test results include the length of the faulty cable segment.

To test the cable connection of an Ethernet interface:

Step	Command
1. Enter system view.	system-view
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>
3. Test the cable connected to the Ethernet interface.	virtual-cable-test

Configuring storm control on an Ethernet interface

Storm control compares broadcast, multicast, and unknown unicast traffic regularly with their respective traffic thresholds on an Ethernet interface. For each type of traffic, storm control provides a lower threshold and a higher threshold.

For management purposes, you can configure the interface to send threshold event traps and log messages when monitored traffic exceeds the upper threshold or drops below the lower threshold from the upper threshold.

When a particular type of traffic exceeds its upper threshold, the interface does either of the following, depending on your configuration:

- Blocks this type of traffic, while forwarding other types of traffic. Even though the interface does not forward the blocked traffic, it still counts the traffic. When the blocked traffic drops below the lower threshold, the port begins to forward the traffic.
- Shuts down automatically. The interface shuts down automatically and stops forwarding any traffic. When the blocked traffic drops below the lower threshold, the port does not forward the traffic. To bring up the interface, use the **undo shutdown** command or disable the storm control function.

Alternatively, you can configure the storm suppression function to control a specific type of traffic. Do not enable them both on an Ethernet interface at the same time because the storm suppression and storm control functions are mutually exclusive. For example, with an unknown unicast suppression threshold set on an Ethernet interface, do not enable storm control for unknown unicast traffic on the interface. For more information about storm suppression, see "[Configuring storm suppression](#)."

Configuration restrictions and guidelines

- For network stability, use the default or set a higher traffic polling interval.
- Storm control uses a complete polling cycle to collect traffic data, and analyzes the data in the next cycle. It takes a port at least one polling interval and at most two polling intervals to take a storm control action.
- The storm control function allows you to set the upper and lower thresholds for all three types of packets respectively on the same interface.

Configuration procedure

To configure storm control on an Ethernet interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the traffic polling interval of the storm control module.	storm-constrain interval <i>seconds</i>	Optional. 10 seconds by default.
3. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Enable storm control, and set the lower and upper thresholds for broadcast, multicast, or unknown unicast traffic.	storm-constrain { broadcast multicast unicast } { pps kpps ratio } <i>max-pps-values</i> <i>min-pps-values</i>	Disabled by default.
5. Set the control action to take when monitored traffic exceeds the upper threshold.	storm-constrain control { block shutdown }	Optional. Disabled by default.
6. Enable the interface to send storm control threshold event traps..	storm-constrain enable trap	Optional. By default, the interface sends traps when monitored traffic exceeds the upper threshold or drops below the lower threshold from the upper threshold.
7. Enable the interface to log storm control threshold events..	storm-constrain enable log	Optional. By default, the interface outputs log messages when monitored traffic exceeds the upper threshold or drops below the lower threshold from the upper threshold.

Setting the MTU for a Layer 3 Ethernet interface (available only on the 5500 EI)

The maximum transmission unit (MTU) affects the fragmentation and re-assembly of IP packets.

To set the MTU for a Layer 3 Ethernet interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Set the MTU.	mtu <i>size</i>	Optional. 1500 bytes by default.

Displaying and maintaining an Ethernet interface

Task	Command	Remarks
Display Ethernet interface information.	display interface [<i>interface-type</i>] brief [down] [{ begin exclude include } <i>regular-expression</i>] display interface <i>interface-type</i> <i>interface-number</i> [brief] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display traffic statistics for the specified interfaces.	display counters { inbound outbound } interface [<i>interface-type</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display traffic rate statistics over the last sampling interval.	display counters rate { inbound outbound } interface [<i>interface-type</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about discarded packets on the specified interfaces.	display packet-drop interface [<i>interface-type</i> [<i>interface-number</i>]] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display summary information about discarded packets on all interfaces.	display packet-drop summary [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the combo interfaces and the fiber and copper combo ports.	display port combo [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about a manual port group or all manual port groups.	display port-group manual [all name <i>port-group-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about the loopback function.	display loopback-detection [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear the interface statistics.	reset counters interface [<i>interface-type</i> [<i>interface-number</i>]]	Available in user view
Clear the statistics of discarded packets on the specified interfaces.	reset packet-drop interface [<i>interface-type</i> [<i>interface-number</i>]]	Available in user view

Configuring loopback and null interfaces

Configuring a loopback interface

Introduction to the loopback interface

A loopback interface is a software-only virtual interface. It delivers the following benefits:

- The physical layer state and link-layer protocols of a loopback interface are always up unless the loopback interface is manually shut down.
- To save IP address resources, you can assign an IP address with an all-F mask to a loopback interface. When you assign an IPv4 address whose mask is not 32-bit, the system automatically changes the mask into a 32-bit mask. When you assign an IPv6 address whose mask is not 128-bit, the system automatically changes the mask into a 128-bit mask.
- You can enable routing protocols on a loopback interface, and a loopback interface can send and receive routing protocol packets.

Because of the benefits mentioned above, loopback interfaces are widely used in the following scenarios:

- You can configure a loopback interface address as the source address of the IP packets that the device generates. Because loopback interface addresses are stable unicast addresses, they are usually used as device identifications. When you configure a rule on an authentication or security server to permit or deny packets generated by a device, you can simplify the rule by configuring it to permit or deny packets that carry the loopback interface address identifying the device. When you use a loopback interface address as the source address of IP packets, be sure to perform any necessary routing configuration to make sure that the route from the loopback interface to the peer is reachable. All data packets sent to the loopback interface are treated as packets sent to the device itself, so the device does not forward these packets.
- Because a loopback interface is always up, it can be used in dynamic routing protocols. For example, if you do not configure a router ID for a dynamic routing protocol, the highest loopback interface IP address is selected as the router ID. In BGP, to prevent physical port failure from interrupting BGP sessions, you can use a loopback interface as the source interface of BGP packets.

Configuration procedure

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a loopback interface and enter loopback interface view.	interface loopback <i>interface-number</i>	N/A
3. Set the interface description.	description <i>text</i>	Optional By default, the description of a loopback interface is <i>interface name</i> Interface.

Step	Command	Remarks
4. Shut down the loopback interface.	shutdown	Optional By default, a loopback interface is up.
5. Restore the default settings for the loopback interface.	default	Optional

NOTE:

You can configure settings such as IP addresses and IP routes on loopback interfaces. For more information, see *Layer 3—IP Services Configuration Guide* and *Layer 3—IP Routing Configuration Guide*.

Configuring a null interface

Introduction to the null interface

A null interface is a completely software-based logical interface, and is always up. However, you cannot use it to forward data packets or configure an IP address or link-layer protocol on it. With a null interface specified as the next hop of a static route to a specific network segment, any packets routed to the network segment are dropped. The null interface provides a simpler way to filter packets than ACL. You can filter uninteresting traffic by transmitting it to a null interface instead of applying an ACL.

For example, by executing the **ip route-static 92.101.0.0 255.255.0.0 null 0** command (which configures a static route that leads to null interface 0), you can have all the packets destined to the network segment 92.101.0.0/16 discarded.

Only one null interface, Null 0, is supported on your switch. You cannot remove or create a null interface.

Configuration procedure

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter null interface view.	interface null 0	The Null 0 interface is the default null interface on your switch. It cannot be manually created or removed.
3. Set the interface description.	description text	Optional By default, the description of a null interface is <i>interface name</i> Interface.
4. Restore the default settings for the null interface.	default	Optional

Displaying and maintaining loopback and null interfaces

Task	Command	Remarks
Display information about loopback interfaces.	display interface [loopback] [brief [down]] [{ begin exclude include } <i>regular-expression</i>] display interface loopback <i>interface-number</i> [brief] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about the null interface.	display interface [null] [brief [down]] [{ begin exclude include } <i>regular-expression</i>] display interface null 0 [brief] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear the statistics on a loopback interface.	reset counters interface [loopback] [<i>interface-number</i>]	Available in user view
Clear the statistics on the null interface.	reset counters interface [null [0]]	Available in user view

Bulk configuring interfaces

You can enter interface range view to bulk configure multiple interfaces with the same feature instead of configuring them one by one. For example, you can perform the **shutdown** command in interface range view to shut down a range of interfaces.

Failure of applying a command on one member interface does not affect the application of the command on the other member interfaces. If applying a command on one member interface fails, the system displays an error message and continues with the next member interface.

Configuration guidelines

When you bulk configure interfaces in interface range view, follow these restrictions and guidelines:

- In interface range view, only the commands supported by the first interface are available.
- Do not assign an aggregate interface and any of its member interfaces to an interface range at the same time. Some commands, after being executed on both an aggregate interface and its member interfaces, can break up the aggregation.
- No limit is set on the maximum number of interfaces in an interface range. The more interfaces in an interface range, the longer the command execution time.

Configuration procedure

To bulk configure interfaces:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface range view.	Approach 1: interface range { <i>interface-type</i> <i>interface-number</i> [to <i>interface-type</i> <i>interface-number</i>] } &<1-5> Approach 2: interface range name <i>name</i> [interface { <i>interface-type</i> <i>interface-number</i> [to <i>interface-type</i> <i>interface-number</i>] } &<1-5>]	Use either approach. In approach 2, you assign a name to an interface range and can specify this name rather than the interface range to enter the interface range view.
3. Display commands available for the first interface in the interface range.	Enter ? at the interface range prompt.	Optional.
4. Perform available commands to configure the interfaces.	Available commands vary by interface.	N/A
5. Verify the configuration.	display this	Optional.

Configuring the MAC address table

This feature covers only the unicast MAC address table. For information about configuring static multicast MAC address table entries for IGMP snooping and MLD snooping, see *IP Multicast Configuration Guide*.

The MAC address table can contain only Layer 2 Ethernet ports and Layer 2 aggregate interfaces.

The MAC address table configuration tasks are all optional and can be performed in any order.

Overview

To reduce single-destination packet flooding in a switched LAN, an Ethernet device uses a MAC address table for forwarding frames through unicast instead of broadcast. This table describes from which port a MAC address (or host) can be reached. When forwarding a single-destination frame, the device first looks up the MAC address of the frame in the MAC address table for a match. If the switch finds an entry, it forwards the frame out of the outgoing port in the entry. If the switch does not find an entry, it floods the frame out of all but the incoming port.

How a MAC address table entry is created

The switch automatically obtains entries in the MAC address table, or you can add them manually.

MAC address learning

The device can automatically populate its MAC address table by obtaining the source MAC addresses (called "MAC address learning") of incoming frames on each port.

When a frame arrives at a port, Port A, for example, the device performs the following tasks:

1. Verifies the source MAC address (for example, MAC-SOURCE) of the frame.
2. Looks up the source MAC address in the MAC address table.
 - If an entry is found, the device updates the entry.
 - If no entry is found, the device adds an entry for MAC-SOURCE and Port A.
3. After obtaining this source MAC address, when the device receives a frame destined for MAC-SOURCE, the device finds the MAC-SOURCE entry in the MAC address table and forwards the frame out of Port A.

The device performs this learning process each time it receives a frame from an unknown source MAC address, until the MAC address table is fully populated.

Manually configuring MAC address entries

With dynamic MAC address learning, a device does not distinguish between illegitimate and legitimate frames, which can invite security hazards. For example, when a hacker sends frames with a forged source MAC address to a port different from the one to which the real MAC address is connected, the device creates an entry for the forged MAC address, and forwards frames destined for the legal user to the hacker instead.

To improve port security, you can bind specific user devices to the port by manually adding MAC address entries to the MAC address table of the switch.

Types of MAC address table entries

A MAC address table can contain the following types of entries:

- **Static entries**—Manually added and never age out.
- **Dynamic entries**—Manually added or dynamically obtained, and might age out.
- **Blackhole entries**—**Manually** configured and never age out. Blackhole entries are configured for filtering out frames with specific source or destination MAC addresses. For example, to block all packets destined for a specific user for security concerns, you can configure the MAC address of this user as a blackhole MAC address entry.

A static or blackhole MAC address entry can overwrite a dynamic MAC address entry, but not vice versa.

To adapt to network changes and prevent inactive entries from occupying table space, an aging mechanism is adopted for dynamic MAC address entries. Each time a dynamic MAC address entry is obtained or created, an aging time starts. If the entry has not updated when the aging timer expires, the device deletes the entry. If the entry has updated before the aging timer expires, the aging timer restarts.

MAC address table-based frame forwarding

When forwarding a frame, the device adopts the following forwarding modes based on the MAC address table:

- **Unicast mode:** If an entry is available for the destination MAC address, the device forwards the frame out of the outgoing interface indicated by the MAC address table entry.
- **Broadcast mode:** If the device receives a frame with the destination address as all-ones, or no entry is available for the destination MAC address, the device broadcasts the frame to all the interfaces except the receiving interface.

Configuring static, dynamic, and blackhole MAC address table entries

To prevent MAC address spoofing attacks and improve port security, you can manually add MAC address table entries to bind ports with MAC addresses. You can also configure blackhole MAC address entries to filter out packets with certain source or destination MAC addresses.

Adding or modifying a static, dynamic, or blackhole MAC address table entry in system view

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Add or modify a dynamic or static MAC address entry.	mac-address { dynamic static } <i>mac-address interface interface-type interface-number vlan vlan-id</i>	Use either command.
3. Add or modify a blackhole MAC address entry.	mac-address blackhole mac-address vlan <i>vlan-id</i>	Make sure that you have created the VLAN and assigned the interface to the VLAN.

Adding or modifying a static or dynamic MAC address table entry in interface view

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Add or modify a static or dynamic MAC address entry.	mac-address { dynamic static } <i>mac-address</i> vlan <i>vlan-id</i>	Make sure that you have created the VLAN and assigned the interface to the VLAN.

Disabling MAC address learning

Sometimes, you might need to disable MAC address learning to prevent the MAC address table from being saturated, for example, when your device is being attacked by a large amount of packets with different source MAC addresses.

When MAC address learning is disabled, the learned MAC addresses remain valid until they age out.

Disabling global MAC address learning

Disabling global MAC address learning disables the learning function on all ports.

To disable MAC address learning:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Disable global MAC address learning.	mac-address mac-learning disable	Enabled by default.

Disabling MAC address learning on ports

After enabling global MAC address learning, you can disable the function on a single port, or on all ports in a port group as needed.

To disable MAC address learning on an interface or a port group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable global MAC address learning.	undo mac-address mac-learning disable	Optional Enabled by default.

Step	Command	Remarks
3. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command. Settings in Layer 2 Ethernet interface view or Layer 2 aggregate interface view take effect on the interface only. Settings in port group view take effect on all member ports in the port group.
4. Disable MAC address learning on the interface or all ports in the port group.	mac-address mac-learning disable	Enabled by default. For more information about configuring a port group, see " Configuring Ethernet interfaces. "

Disabling MAC address learning on a VLAN (available only on the 5500 EI)

You can disable MAC address learning on a per-VLAN basis.

To disable MAC address learning on a VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable global MAC address learning.	undo mac-address mac-learning disable	Optional Enabled by default.
3. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
4. Disable MAC address learning on the VLAN.	mac-address mac-learning disable	Enabled by default.

Configuring the aging timer for dynamic MAC address entries

The MAC address table uses an aging timer for dynamic MAC address entries for security and efficient use of table space. If a dynamic MAC address entry has failed to update before the aging timer expires, the device deletes that entry. This aging mechanism ensures that the MAC address table can quickly update to accommodate the latest network changes.

Set the aging timer appropriately. Too long an aging interval might cause the MAC address table to retain outdated entries, exhaust the MAC address table resources, and fail to update its entries to accommodate the latest network changes. Too short an interval might result in removal of valid entries, causing unnecessary flooding, which might affect device performance.

To configure the aging timer for dynamic MAC address entries:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Configure the aging timer for dynamic MAC address entries.	mac-address timer { aging <i>seconds</i> no-aging }	Optional 300 seconds by default. The no-aging keyword disables the aging timer.

You can reduce flooding on a stable network by disabling the aging timer to prevent dynamic entries from unnecessarily aging out. By reducing flooding, you improve not only network performance, but also security, because you reduce the chances that a data packet will reach unintended destinations.

Configuring the MAC learning limit on ports

To prevent the MAC address table from getting too large, you can limit the number of MAC addresses that a port can learn.

To configure the MAC learning limit on a Layer 2 Ethernet interface or all ports in a port group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command. Settings in Layer 2 Ethernet interface view take effect on the interface only. Settings in port group view take effect on all member ports in the port group.
3. Configure the MAC learning limit on the interface or port group.	mac-address max-mac-count <i>count</i>	No MAC learning limit is configured by default. Layer 2 aggregate interfaces do not support this command.

NOTE:

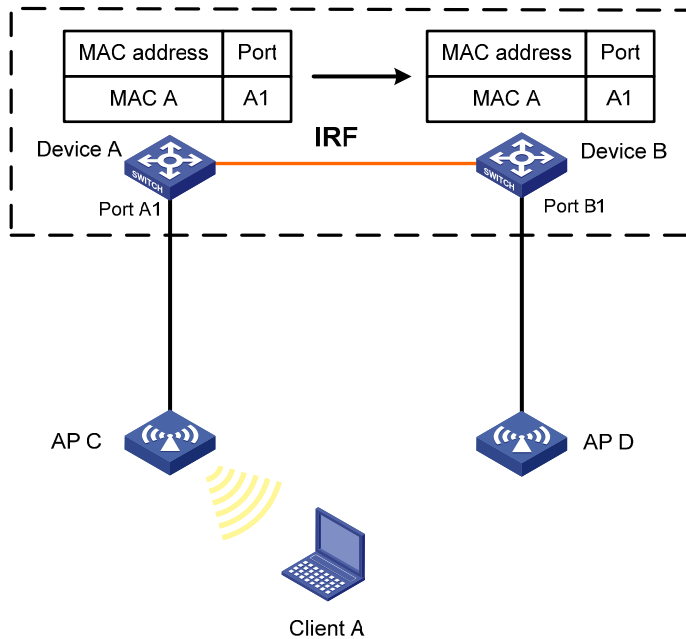
Do not configure the MAC learning limit on any member ports of an aggregation group. Otherwise, the member ports cannot be selected.

Enabling MAC address roaming

After you enable MAC address roaming on an IRF fabric, each member switch advertises learned MAC addresses to other member switches.

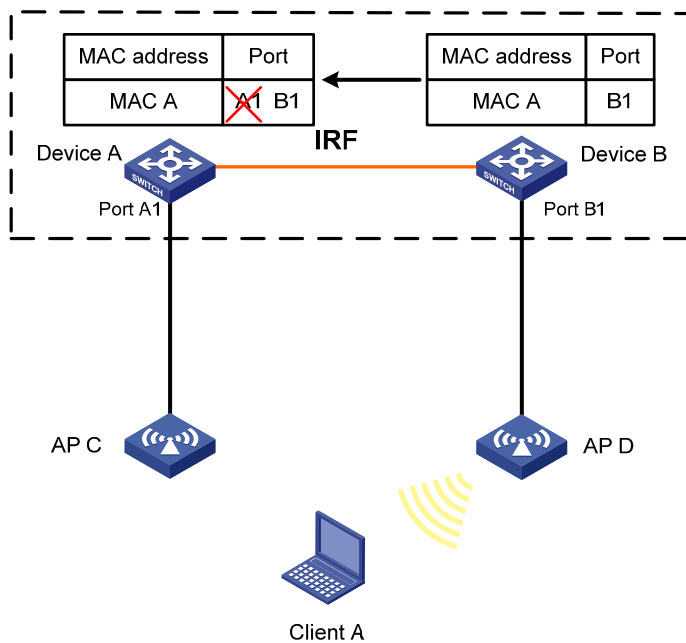
As shown in [Figure 4](#), Device A and Device B form an IRF fabric enabled with MAC address roaming. They connect to AP C and AP D, respectively. When Client A associates with AP C, Device A learns the MAC address of Client A and advertises it to the member switch Device B.

Figure 4 MAC address tables of devices when Client A associates with AP C



If Client A roams to AP D, Device B learns the MAC address of Client A and advertises it to Device A to ensure service continuity for Client A, as shown in Figure 5.

Figure 5 MAC address tables of devices when Client A roams to AP D



To enable MAC address roaming:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable MAC address roaming.	mac-address mac-roaming enable	Disabled by default.

Displaying and maintaining MAC address tables

Task	Command	Remarks
Display MAC address table information.	display mac-address [<i>mac-address</i> [vlan <i>vlan-id</i>]] [[dynamic static] [interface <i>interface-type</i> <i>interface-number</i>] blackhole] [vlan <i>vlan-id</i>] [count]] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the aging timer for dynamic MAC address entries.	display mac-address aging-time [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the system or interface MAC address learning state.	display mac-address mac-learning [<i>interface-type</i> <i>interface-number</i>] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display MAC address statistics.	display mac-address statistics [[{ begin exclude include } <i>regular-expression</i>]	Available in any view

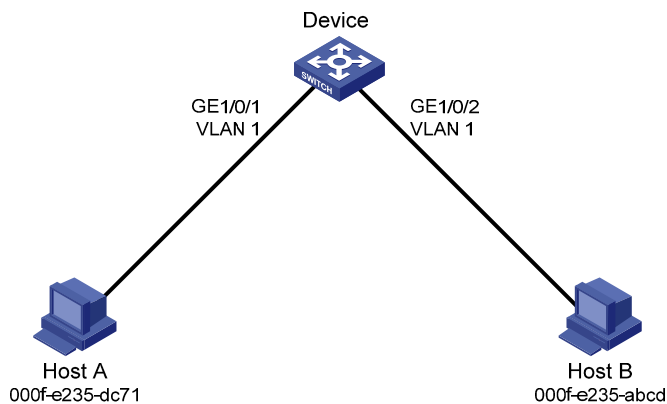
MAC address table configuration example

Network requirements

As shown in [Figure 6](#):

- The MAC address of Host A is 000f-e235-dc71 and belongs to VLAN 1. It is connected to GigabitEthernet 1/0/1 of the device. To prevent MAC address spoofing, add a static entry for the host in the MAC address table of the device.
- The MAC address of Host B is 000f-e235-abcd and belongs to VLAN 1. For security, because this host once behaved suspiciously on the network, add a blackhole MAC address entry for the host MAC address, so all packets destined for the host are dropped.
- Set the aging timer for dynamic MAC address entries to 500 seconds.

Figure 6 Network diagram



Configuration procedure

Add a static MAC address entry.

```
<Sysname> system-view
```

```
[Sysname] mac-address static 000f-e235-dc71 interface gigabitethernet 1/0/1 vlan 1
```

Add a blackhole MAC address entry.

```
[Sysname] mac-address blackhole 000f-e235-abcd vlan 1
```

Set the aging timer for dynamic MAC address entries to 500 seconds.

```
[Sysname] mac-address timer aging 500
```

Display the MAC address entry for port GigabitEthernet 1/0/1.

```
[Sysname] display mac-address interface gigabitethernet 1/0/1
```

MAC ADDR	VLAN ID	STATE	PORT INDEX	AGING TIME(s)
000f-e235-dc71	1	Config static	GigabitEthernet 1/0/1	NOAGED

```
--- 1 mac address(es) found ---
```

Display information about the blackhole MAC address table.

```
[Sysname] display mac-address blackhole
```

MAC ADDR	VLAN ID	STATE	PORT INDEX	AGING TIME(s)
000f-e235-abcd	1	Blackhole	N/A	NOAGED

```
--- 1 mac address(es) found ---
```

View the aging time of dynamic MAC address entries.

```
[Sysname] display mac-address aging-time
```

```
Mac address aging time: 500s
```

Configuring MAC Information

Overview

Introduction to MAC Information

To monitor a network, you must monitor users who are joining and leaving the network. Because a MAC address uniquely identifies a network user, you can monitor users who are joining and leaving a network by monitoring their MAC addresses.

With the MAC Information function, Layer 2 Ethernet ports send Syslog or trap messages to the monitor end in the network when they obtain or delete MAC addresses. By analyzing these messages, the monitor end can monitor users who are accessing the network.

How MAC Information works

When a new MAC address is obtained or an existing MAC address is deleted on a device, the device writes related information about the MAC address to the buffer area used to store user information. When the timer set for sending MAC address monitoring Syslog or trap messages expires, or when the buffer reaches capacity, the device sends the Syslog or trap messages to the monitor end.

The device writes information and sends messages only for the following MAC addresses: automatically learned source MAC addresses, MAC addresses that pass MAC authentication, MAC addresses that pass 802.1X authentication, MAC addresses matching OUI addresses in the voice VLAN feature, and secure MAC addresses. The device does not write information or send messages for blackhole MAC address, static MAC addresses, dynamic MAC addresses that are manually configured, multicast MAC addresses, and local MAC addresses.

For more information about MAC authentication, 802.1X, and secure MAC addresses in port security, see *Security Configuration Guide*. For more information about voice VLAN and OUI addresses, see "Configuring a voice VLAN."

Configuring MAC Information

Enabling MAC Information globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable MAC Information globally.	mac-address information enable	Disabled by default.

Enabling MAC Information on an interface

To enable MAC Information on an Ethernet interface, enable MAC Information globally first.

To enable MAC Information on an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable MAC Information on the interface.	mac-address information enable { added deleted }	Disabled by default.

Configuring MAC Information mode

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure MAC Information mode.	mac-address information mode { syslog trap }	Optional trap by default.

Configuring the interval for sending Syslog or trap messages

To prevent Syslog or trap messages from being sent too frequently, change the interval for sending Syslog or trap messages.

To set the interval for sending Syslog or trap messages:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the interval for sending Syslog or trap messages.	mac-address information interval <i>interval-time</i>	Optional One second by default.

Configuring the MAC Information queue length

To avoid losing user MAC address information, when the buffer that stores user MAC address information reaches capacity, the user MAC address information in the buffer is sent to the monitor end in the network, even if the timer set for sending MAC address monitoring Syslog or trap messages has not expired yet.

To configure the MAC Information queue length:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the MAC Information queue length.	mac-address information queue-length <i>value</i>	Optional 50 by default.

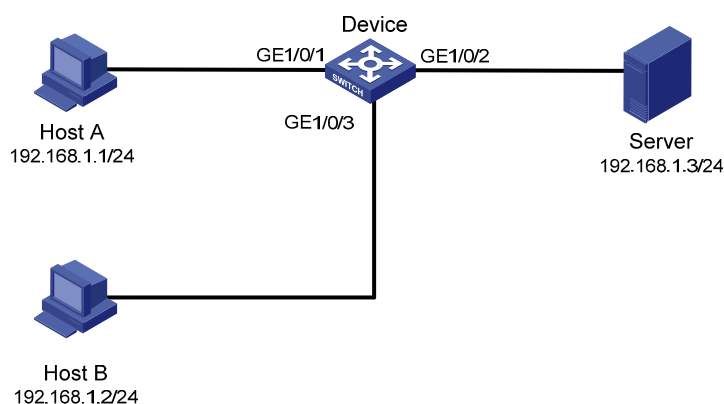
MAC Information configuration example

Network requirements

As shown:

Enable MAC Information on GigabitEthernet 1/0/1 on Device in [Figure 7](#) to send MAC address changes in Syslog messages to Host B through GigabitEthernet 1/0/3. Host B analyzes and displays the Syslog messages.

Figure 7 Network diagram



Configuration procedure

1. Configure Device to send Syslog messages to Host B (see *Network Management and Monitoring Configuration Guide*).
2. Enable MAC Information.

Enable MAC Information on Device.

```
<Device> system-view
```

```
[Device] mac-address information enable
```

Configure MAC Information mode as Syslog.

```
[Device] mac-address information mode syslog
```

Enable MAC Information on GigabitEthernet 1/0/1.

```
[Device] interface gigabitethernet 1/0/1
```

```
[Device-GigabitEthernet1/0/1] mac-address information enable added
```

```
[Device-GigabitEthernet1/0/1] mac-address information enable deleted
```

```
[Device-GigabitEthernet1/0/1] quit
```

Set the MAC Information queue length to 100.

```
[Device] mac-address information queue-length 100
```

Set the interval for sending Syslog or trap messages to 20 seconds.

```
[Device] mac-address information interval 20
```

Configuring Ethernet link aggregation

The 5500 SI Switch Series does not support Layer 3 aggregate interfaces. The Layer 3 aggregate interface configurations in this document apply to only the 5500 EI Switch Series.

Overview

Ethernet link aggregation, or simply link aggregation, combines multiple physical Ethernet ports into one logical link, called an "aggregate link." Link aggregation delivers the following benefits:

- Increases bandwidth beyond the limits of any single link. In an aggregate link, traffic is distributed across the member ports.
- Improves link reliability. The member ports dynamically back up one another. When a member port fails, its traffic is automatically switched to other member ports.

As shown in [Figure 8](#), Device A and Device B are connected by three physical Ethernet links. These physical Ethernet links are combined into an aggregate link, Link Aggregation 1. The bandwidth of this aggregate link is as high as the total bandwidth of the three physical Ethernet links. At the same time, the three Ethernet links back up one another.

Figure 8 Ethernet link aggregation



Basic concepts

Aggregation group, member port, and aggregate interface

Link aggregation is implemented by combining Ethernet interfaces into a link aggregation group. Each link aggregation group has one logical aggregate interface. To an upper layer entity that uses the link aggregation service, a link aggregation group appears to be a single logical link and data traffic is transmitted through the aggregate interface. The rate of an aggregate interface equals the total rate of its member ports in the Selected state, and its duplex mode is the same as the selected member ports. For more information about the states of member ports in an aggregation group, see "[Aggregation states of member ports in an aggregation group](#)."

Aggregate interfaces are one of the following types:

Bridge-aggregation (BAGG) interfaces—Also called Layer 2 aggregate interfaces.

Route-aggregation (RAGG) interfaces—Also called Layer 3 aggregate interfaces.

When you create an aggregate interface, the switch automatically creates an aggregation group of the same type and number as the aggregate interface. For example, when you create interface Bridge-Aggregation 1, Layer 2 aggregation group 1 is automatically created.

You can assign Layer 2 Ethernet interfaces only to a Layer 2 aggregation group, and Layer 3 Ethernet interfaces only to a Layer 3 aggregation group.

Aggregation states of member ports in an aggregation group

A member port in an aggregation group can be in either of the following aggregation states:

- Selected: A Selected port can forward user traffic.
- Unselected: An Unselected port cannot forward user traffic.

Operational key

When aggregating ports, the system automatically assigns each port an operational key based on port information such as port rate and duplex mode. Any change to this information triggers a recalculation of the operational key.

In an aggregation group, all selected member ports are assigned the same operational key.

Configuration classes

Every configuration setting on a port might affect its aggregation state. Port configurations fall into the following classes:

- **Port attribute configurations**—Includes port rate, duplex mode, and link status (up/down). These are the most basic port configurations.
- **Class-two configurations**—A member port can be placed in Selected state only if it has the same class-two configurations as the aggregate interface. Class-two configurations made on an aggregate interface are automatically synchronized to all its member ports. These configurations are retained on the member ports even after the aggregate interface is removed.

Table 2 Class-two configurations

Feature	Considerations
Port isolation	Whether the port has joined an isolation group
QinQ	QinQ enable state (enable/disable), TPID for VLAN tags, outer VLAN tags to be added, inner-to-outer VLAN priority mappings, inner-to-outer VLAN tag mappings, inner VLAN ID substitution mappings
VLAN	Permitted VLAN IDs, PVID, link type (trunk, hybrid, or access), IP subnet-based VLAN configuration, protocol-based VLAN configuration, VLAN tagging mode
MAC address learning	MAC address learning capability

NOTE:

Any class-two configuration change might affect the aggregation state of link aggregation member ports and ongoing traffic. To be sure that you are aware of the risk, the system displays a warning message every time you attempt to change a class-two configuration setting on a member port.

- **Class-one configurations**—Include settings that do not affect the aggregation state of the member port even if they are different from those on the aggregate interface. GVRP and MSTP settings are examples of class-one configurations. The class-one configuration for a member port is effective only when the member port leaves the aggregation group.

Reference port

When setting the aggregation state of the ports in an aggregation group, the system automatically picks a member port as the reference port. A Selected port must have the same port attributes and class-two configurations as the reference port.

LACP

The IEEE 802.3ad Link Aggregation Control Protocol (LACP) enables dynamic aggregation of physical links. It uses link aggregation control protocol data units (LACPDU) for exchanging aggregation information between LACP-enabled devices.

1. LACP functions

The IEEE 802.3ad LACP offers basic LACP functions and extended LACP functions, as described in [Table 3](#).

Table 3 Basic and extended LACP functions

Category	Description
Basic LACP functions	Implemented through the basic LACPDU fields, including the system LACP priority, system MAC address, port aggregation priority, port number, and operational key. Each member port in a LACP-enabled aggregation group exchanges the preceding information with its peer. When a member port receives an LACPDU, it compares the received information with the information received on the other member ports. In this way, the two systems reach an agreement on which ports should be placed in the Selected state.
Extended LACP functions	Implemented by extending the LACPDU with new Type/Length/Value (TLV) fields. This is how the LACP multi-active detection (MAD) mechanism of the Intelligent Resilient Framework (IRF) feature is implemented. The 5500 EI[S] Switch Series can participate in LACP MAD as either an IRF member switch or an intermediate device.

For more information about IRF, member switches, intermediate devices, and the LACP MAD mechanism, see *IRF Configuration Guide*.

2. LACP priorities

LACP priorities have the following types: system LACP priority and port aggregation priority.

Table 4 LACP priorities

Type	Description	Remarks
System LACP priority	Used by two peer devices (or systems) to determine which one is superior in link aggregation. In dynamic link aggregation, the system that has higher system LACP priority sets the Selected state of member ports on its side first, and then the system that has lower priority sets the port state accordingly.	The smaller the priority value, the higher the priority.
Port aggregation priority	Determines the likelihood of a member port to be selected on a system. The higher the port aggregation priority, the higher the likelihood.	

3. LACP timeout interval

The LACP timeout interval specifies how long a member port waits to receive LACPDUs from the peer port. If a local member port fails to receive LACPDUs from the peer within three times the LACP timeout interval, the member port assumes that the peer port has failed. You can configure the LACP timeout interval as either the short timeout interval (1 second) or the long timeout interval (30 seconds).

Link aggregation modes

Link aggregation has the following modes: dynamic and static. Dynamic link aggregation uses LACP and static link aggregation does not. [Table 5](#) compares the two aggregation modes.

Table 5 A comparison between static and dynamic aggregation modes

Aggregation mode	LACP status on member ports	Pros	Cons
Static	Disabled	Aggregation is stable. Peers do not affect the aggregation state of the member ports.	The member ports do not adjust the aggregation state according to that of the peer ports. The administrator must manually maintain link aggregations.
Dynamic	Enabled	The administrator does not need to maintain link aggregations. The peer systems maintain the aggregation state of the member ports automatically.	Aggregation is unstable. The aggregation state of the member ports is susceptible to network changes.

The following points apply to a dynamic link aggregation group:

- A Selected port can receive and send LACPDUs.
- An Unselected port can receive and send LACPDUs only if it is up and has the same class-two configurations as the aggregate interface.

Aggregating links in static mode

LACP is disabled on the member ports in a static aggregation group. You must manually maintain the aggregation state of the member ports.

The static link aggregation process comprises:

- [Selecting a reference port](#)
- [Setting the aggregation state of each member port](#)

Selecting a reference port

The system selects a reference port from the member ports that are:

- Are in the up state and have
- Have the same class-two configurations as the aggregate interface.

The candidate ports are sorted by aggregation priority, duplex, and speed in the following order:

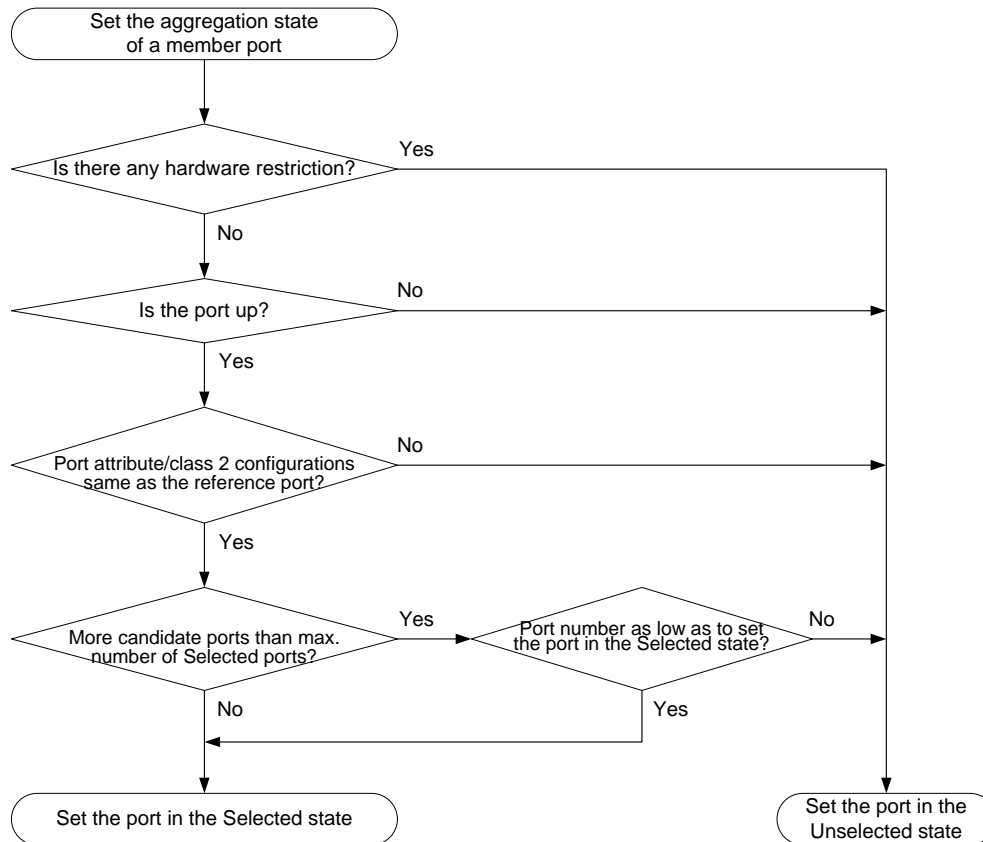
- Lowest aggregation priority value
- Full duplex/high speed
- Full duplex/low speed
- Half duplex/high speed
- Half duplex/low speed

The one at the top is selected as the reference port. If two ports have the same aggregation priority, duplex mode, and speed, the one with the lower port number wins.

Setting the aggregation state of each member port

After selecting the reference port, the static aggregation group sets the aggregation state of each member port, as shown in [Figure 9](#). After the static aggregation group has reached the limit on Selected ports, any port assigned to the group is placed in Unselected state to avoid traffic interruption on the current Selected ports.

Figure 9 Setting the aggregation state of a member port in a static aggregation group



Aggregating links in dynamic mode

LACP is automatically enabled on all member ports in a dynamic aggregation group. The protocol automatically maintains the aggregation state of ports.

The dynamic link aggregation process comprises:

- Selecting a reference port
- Setting the aggregation state of each member port

Selecting a reference port

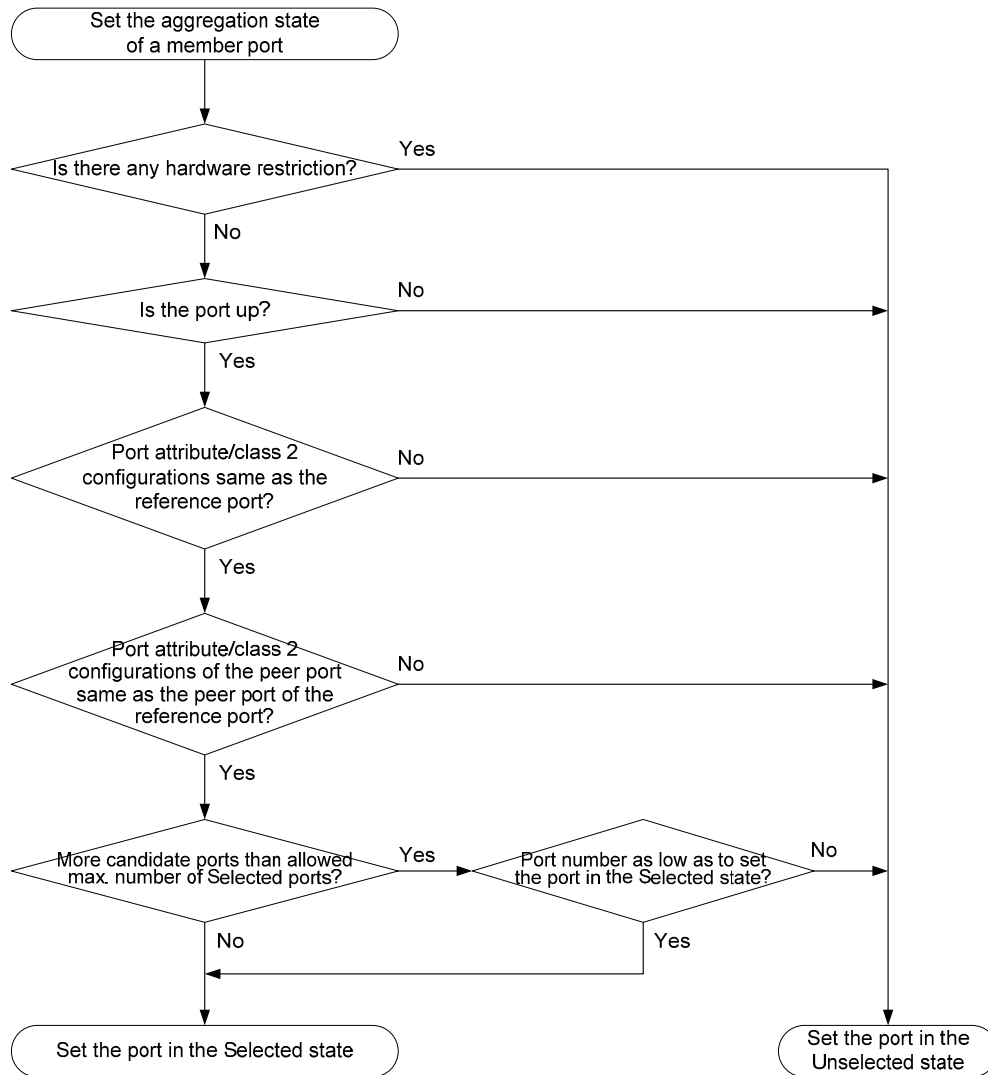
The local system (the actor) and the remote system (the partner) negotiate a reference port using the following workflow:

1. The systems compare the system ID (which comprises the system LACP priority and the system MAC address). The system with the lower LACP priority value wins. If they are the same, the systems compare the system MAC addresses. The system with the lower MAC address wins.
2. The system with the smaller system ID selects the port with the smallest port ID as the reference port. A port ID comprises a port aggregation priority and a port number. The port with the lower aggregation priority value wins. If two ports have the same aggregation priority, the system compares their port numbers. The port with the smaller port number wins.

Setting the aggregation state of each member port

After the reference port is selected, the system with the lower system ID sets the state of each member port in the dynamic aggregation group on its side.

Figure 10 Setting the state of a member port in a dynamic aggregation group



Meanwhile, the system with the higher system ID, which has identified the aggregation state changes on the remote system, sets the aggregation state of local member ports as the same as their peer ports.

A dynamic link aggregation group preferably sets full-duplex ports as the Selected ports, and will set one, and only one, half-duplex port as a Selected port when none of the full-duplex ports can be selected or only half-duplex ports exist in the group.

When the aggregation state of a member port changes, the aggregation state of its peer port also changes.

After the Selected port limit has been reached, a port assigned to the dynamic aggregation group is placed in Selected state if it is more eligible for being selected than a current member port.

The port assigned to the dynamic aggregation group after the Selected port limit has been reached is placed in Selected state if it is more eligible for being selected than a current member port.

Load-sharing criteria for link aggregation groups

In a link aggregation group, traffic can be load-shared across the selected member ports based on a set of criteria, depending on your configuration.

You can choose one of the following criteria or any combination for load sharing:

- MAC addresses
- Service port numbers
- Ingress ports
- IP addresses

Alternatively, you can let the system automatically choose link-aggregation load-sharing criteria based on packet types (Layer 2, IPv4, or IPv6 for example)

Configuration restrictions and guidelines

Follow these guidelines when you configure a link aggregation group:

- To ensure stable aggregation state and service continuity, do not change port attributes or class-two configurations on any member port. If you must, make sure you understand its impact on the live network. Any port attribute or class-two configuration change might affect the aggregation state of link aggregation member ports and ongoing traffic.

Avoid assigning ports to a static aggregation group that has reached the limit on Selected ports. These ports will be placed in Unselected state to avoid traffic interruption on the current Selected ports. However, a device reboot can cause the aggregation state of member ports to change.

Ethernet link aggregation configuration task list

Complete the following tasks to configure Ethernet link aggregation:

Task	Remarks	
Configuring an aggregation group	Configuring a static aggregation group	
	Configuring a dynamic aggregation group	
Configuring an aggregate interface	Select either task	
	Configuring the description of an aggregate interface	Optional
	Configuring the MTU of a Layer 3 aggregate interface	Optional
	Enabling link state traps for an aggregate interface	Optional
	Limiting the number of Selected ports for an aggregation group	Optional
	Shutting down an aggregate interface	Optional
Configuring load sharing for link aggregation groups	Restoring the default settings for an aggregate interface	Optional
	Configuring load-sharing criteria for link aggregation groups	Optional
	Enabling local-first load sharing for link aggregation	Optional
Enabling link-aggregation traffic redirection	Optional	

Configuring an aggregation group

You can choose to create a Layer 2 or Layer 3 link aggregation group depending on the ports to be aggregated on the 5500 EI switch:

- To aggregate Layer 2 Ethernet interfaces, create a Layer 2 link aggregation group.

- To aggregate Layer 3 Ethernet interfaces, create a Layer 3 link aggregation group.

Configuration guidelines

- You cannot assign a port to a Layer 2 aggregation group if any of the features listed in [Table 6](#) is configured on the port.

Table 6 Features incompatible with Layer 2 aggregation groups

Feature	Reference
RRPP	RRPP in <i>High Availability Configuration Guide</i>
MAC authentication	MAC authentication in <i>Security Configuration Guide</i>
Port security	Port security in <i>Security Configuration Guide</i>
IP source guard	IP source guard in <i>Security Configuration Guide</i>
802.1X	802.1X in <i>Security Configuration Guide</i>
Ports specified as source interfaces in portal-free rules	Portal in <i>Security Configuration Guide</i>

- You cannot assign a port to a Layer 3 aggregation group if any of the features listed in [Table 7](#) is configured on the port.

Table 7 Interfaces that cannot be assigned to a Layer 3 aggregation group

Interface type	Reference
Interfaces configured with IP addresses	IP addressing in <i>Layer 3—IP Services Configuration Guide</i>
Interfaces configured as DHCP/BOOTP clients	DHCP in <i>Layer 3—IP Services Configuration Guide</i>
VRRP	VRRP in <i>High Availability Configuration Guide</i>
Portal	Portal in <i>Security Configuration Guide</i>

- Removing an aggregate interface also removes the corresponding aggregation group. At the same time, all member ports leave the aggregation group.

Configuring a static aggregation group

To guarantee a successful static aggregation, make sure that the ports at both ends of each link are in the same aggregation state.

Configuring a Layer 2 static aggregation group

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a Layer 2 aggregate interface and enter Layer 2 aggregate interface view.	interface bridge-aggregation <i>interface-number</i>	When you create a Layer 2 aggregate interface, the system automatically creates a Layer 2 static aggregation group numbered the same.
3. Exit to system view.	quit	N/A

Step	Command	Remarks
4. Assign an Ethernet interface to the aggregation group.	<ul style="list-style-type: none"> a. Enter Layer 2 Ethernet interface view: interface <i>interface-type interface-number</i> b. Assign the Ethernet interface to the aggregation group: port link-aggregation group <i>number</i> 	Repeat this step to assign more Layer 2 Ethernet interfaces to the aggregation group.
5. Assign the port an aggregation priority.	link-aggregation port-priority <i>port-priority</i>	<p>Optional</p> <p>By default, the aggregation priority of a port is 32768.</p> <p>Changing the aggregation priority of a port might affect the aggregation state of the ports in the static aggregation group.</p>

Configuring a Layer 3 static aggregation group (available only on the 5500 EI)

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a Layer 3 aggregate interface and enter Layer 3 aggregate interface view.	interface route-aggregation <i>interface-number</i>	When you create a Layer 3 aggregate interface, the system automatically creates a Layer 3 static aggregation group numbered the same.
3. Exit to system view.	quit	N/A
4. Assign an Ethernet interface to the aggregation group.	<ul style="list-style-type: none"> a. Enter Layer 3 Ethernet interface view: interface <i>interface-type interface-number</i> b. Assign the Ethernet interface to the aggregation group: port link-aggregation group <i>number</i> 	Repeat this step to assign more Layer 3 Ethernet interfaces to the aggregation group.
5. Assign the port an aggregation priority.	link-aggregation port-priority <i>port-priority</i>	<p>Optional</p> <p>By default, the aggregation priority of a port is 32768.</p> <p>Changing the aggregation priority of a port might affect the aggregation state of the ports in the static aggregation group.</p>

Configuring a dynamic aggregation group

To guarantee a successful dynamic aggregation, be sure that the peer ports of the ports aggregated at one end are also aggregated. The two ends can automatically negotiate the aggregation state of each member port.

Configuring a Layer 2 dynamic aggregation group

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the system LACP priority.	lacp system-priority <i>system-priority</i>	Optional By default, the system LACP priority is 32768. Changing the system LACP priority might affect the aggregation state of the ports in a dynamic aggregation group.
3. Create a Layer 2 aggregate interface and enter Layer 2 aggregate interface view.	interface bridge-aggregation <i>interface-number</i>	When you create a Layer 2 aggregate interface, the system automatically creates a Layer 2 static aggregation group numbered the same.
4. Configure the aggregation group to operate in dynamic aggregation mode.	link-aggregation mode dynamic	By default, an aggregation group operates in static aggregation mode.
5. Exit to system view.	quit	N/A
6. Assign an Ethernet interface to the aggregation group.	<ol style="list-style-type: none"> a. Enter Layer 2 Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i> b. Assign the Ethernet interface to the aggregation group: port link-aggregation group <i>number</i> 	Repeat this step to assign more Layer 2 Ethernet interfaces to the aggregation group.
7. Assign the port an aggregation priority.	link-aggregation port-priority <i>port-priority</i>	Optional By default, the aggregation priority of a port is 32768. Changing the aggregation priority of a port might affect the aggregation state of the ports in the dynamic aggregation group.
8. Set the LACP timeout interval on the port to the short timeout interval (1 second).	lacp period short	Optional By default, the LACP timeout interval on a port is the long timeout interval (30 seconds).

Configuring a Layer 3 dynamic aggregation group (available only on the 5500 EI)

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Set the system LACP priority.	lacp system-priority <i>system-priority</i>	Optional By default, the system LACP priority is 32768. Changing the system LACP priority might affect the aggregation state of the ports in the dynamic aggregation group.
3. Create a Layer 3 aggregate interface and enter Layer 3 aggregate interface view.	interface route-aggregation <i>interface-number</i>	When you create a Layer 3 aggregate interface, the system automatically creates a Layer 3 static aggregation group numbered the same.
4. Configure the aggregation group to operate in dynamic aggregation mode.	link-aggregation mode dynamic	By default, an aggregation group operates in static aggregation mode.
5. Exit to system view.	quit	N/A
6. Assign an Ethernet interface to the aggregation group.	<ul style="list-style-type: none"> a. Enter Layer 3 Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i> b. Assign the Ethernet interface to the aggregation group: port link-aggregation group <i>number</i> 	Repeat this step to assign more Layer 3 Ethernet interfaces to the aggregation group.
7. Assign the port an aggregation priority.	link-aggregation port-priority <i>port-priority</i>	Optional By default, the aggregation priority of a port is 32768. Changing the aggregation priority of a port might affect the aggregation state of ports in the dynamic aggregation group.
8. Set the LACP timeout interval on the port to the short timeout interval (1 second).	lacp period short	Optional By default, the LACP timeout interval on a port is the long timeout interval (30 seconds).

Configuring an aggregate interface

Most of the configurations that can be performed on Layer 2 or Layer 3 Ethernet interfaces can also be performed on Layer 2 or Layer 3 aggregate interfaces.

Configuring the description of an aggregate interface

You can configure the description of an aggregate interface for administration purposes such as describing the purpose of the interface.

To configure the description of an aggregate interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter aggregate interface view.	<ul style="list-style-type: none"> Enter Layer 2 aggregate interface view: interface bridge-aggregation <i>interface-number</i> Enter Layer 3 aggregate interface view: interface route-aggregation <i>interface-number</i> 	Use either command.
3. Configure the description of the aggregate interface.	description <i>text</i>	Optional By default, the description of an interface is in the format of <i>interface-name</i> Interface , such as Bridge-Aggregation1 Interface .

Configuring the MTU of a Layer 3 aggregate interface (available only on the 5500 EI)

The MTU of an interface affects IP packets fragmentation and reassembly on the interface.

To change the MTU of a Layer 3 aggregate interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 3 aggregate interface view.	interface route-aggregation <i>interface-number</i>	N/A
3. Configure the MTU of the Layer 3 aggregate interface.	mtu <i>size</i>	Optional 1500 bytes by default.

Enabling link state traps for an aggregate interface

You can configure an aggregate interface to generate linkUp trap messages when its link goes up and linkDown trap messages when its link goes down. For more information, see *Network Management and Monitoring Configuration Guide*.

To enable link state traps on an aggregate interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable the trap function globally.	snmp-agent trap enable [standard [linkdown linkup] *]	Optional By default, link state trapping is enabled globally and on all interfaces.

Step	Command	Remarks
3. Enter aggregate interface view.	<ul style="list-style-type: none"> Enter Layer 2 aggregate interface view: interface bridge-aggregation <i>interface-number</i> Enter Layer 3 aggregate interface view: interface route-aggregation <i>interface-number</i> 	Use either command.
4. Enable link state traps for the aggregate interface.	enable snmp trap updown	Optional Enabled by default.

Limiting the number of Selected ports for an aggregation group

The bandwidth of an aggregate link increases along with the number of selected member ports. To avoid congestion caused by insufficient Selected ports on an aggregate link, you can set the minimum number of Selected ports required for bringing up the specific aggregate interface.

This minimum threshold setting affects the aggregation state of both aggregation member ports and the aggregate interface in the following ways:

- All member ports change to the Unselected state and the link of the aggregate interface goes down, when the number of member ports eligible for being selected is smaller than the minimum threshold.
- When the minimum threshold is reached, the eligible member ports change to the Selected state, and the link of the aggregate interface goes up.

By default, the maximum number of Selected ports allowed in an aggregation group depends on the hardware capabilities of the member ports. After you manually configure the maximum number of Selected ports in an aggregation group, the maximum number of Selected ports allowed in the aggregation group is the lower value of the two upper limits.

You can configure redundancy between two ports using the following guideline: Assign two ports to an aggregation group, and configure the maximum number of Selected ports allowed in the aggregation group as 1. In this way, only one Selected port is allowed in the aggregation group at any point in time, while the Unselected port serves as a backup port.

Configuration guidelines

Follow these guidelines when you configure the port threshold settings:

- If you set a minimum threshold for a static aggregation group, also make the same setting for its peer aggregation group to guarantee correct aggregation.
- Make sure the two link aggregation ends have the same minimum and maximum numbers of selected ports.

Make sure you understand the following impacts of the port threshold settings:

- Configuring the minimum number of Selected ports required to bring up an aggregation group may cause all the member ports in the aggregation group to become unselected.
- Configuring the maximum number of Selected ports in an aggregation group may cause some of the selected member ports in the aggregation group to become unselected.

To limit the number of Selected ports for an aggregation group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter aggregate interface view.	<ul style="list-style-type: none"> Enter Layer 2 aggregate interface view: interface bridge-aggregation <i>interface-number</i> Enter Layer 3 aggregate interface view: interface route-aggregation <i>interface-number</i> 	Use either command.
3. Set the minimum number of Selected ports for the aggregation group.	link-aggregation selected-port minimum <i>number</i>	Not specified by default.
4. Set the maximum number of Selected ports for the aggregation group.	link-aggregation selected-port maximum <i>number</i>	By default, the maximum number of Selected ports allowed in an aggregation group depends on only the hardware capabilities of the member ports.

Shutting down an aggregate interface

Shutting down or bringing up an aggregate interface affects the aggregation state and link state of ports in the corresponding aggregation group in the following ways:

- When an aggregate interface is shut down, all Selected ports in the corresponding aggregation group become unselected and their link state becomes down.
- When an aggregate interface is brought up, the aggregation state of ports in the corresponding aggregation group is recalculated and their link state becomes up.

To shut down an aggregate interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter aggregate interface view.	<ul style="list-style-type: none"> Enter Layer 2 aggregate interface view: interface bridge-aggregation <i>interface-number</i> Enter Layer 3 aggregate interface view: interface route-aggregation <i>interface-number</i> 	Use either command.
3. Shut down the aggregate interface.	shutdown	By default, aggregate interfaces are up.

Restoring the default settings for an aggregate interface

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter aggregate interface view.	<ul style="list-style-type: none">Enter Layer 2 aggregate interface view: interface bridge-aggregation <i>interface-number</i>Enter Layer 3 aggregate interface view: interface route-aggregation <i>interface-number</i>	Use either command.
3. Restore the default settings for the aggregate interface.	default	N/A

Configuring load sharing for link aggregation groups

Configuring load-sharing criteria for link aggregation groups

You can determine how traffic is load-shared in a link aggregation group by configuring load-sharing criteria. The criteria can be MAC addresses, service port numbers, ingress ports, or IP addresses, or any combination. You can also let the system automatically choose link-aggregation load-sharing criteria based on packet types (Layer 2, IPv4, or IPv6, for example).

You can configure global or group-specific load-sharing criteria. A link aggregation group preferentially uses the group-specific load-sharing criteria. If no group-specific load-sharing criteria are available, the group uses the global load-sharing criteria.

NOTE:

The load sharing criteria configuration applies to only unicast packets, and can change the load sharing criteria for unicast packets. Broadcast packets and multicast packets always use the default load sharing criteria.

Configuring the global link-aggregation load-sharing criteria

To configure the global link-aggregation load-sharing criteria:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the global link-aggregation load-sharing criteria.	link-aggregation load-sharing mode { { destination-ip destination-mac destination-port ingress-port source-ip source-mac source-port } }	By default, the system selects the global load sharing criteria according to the packet type.

In system view, the switch supports the following load-sharing criteria and combinations:

- Source IP address
- Destination IP address
- Source MAC address
- Destination MAC address
- Source IP address and destination IP address
- Source IP address and source port
- Destination IP address and destination port
- Any combination of incoming port, source MAC address, and destination MAC address

Configuring group-specific load-sharing criteria

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 aggregate interface view.	interface bridge-aggregation <i>interface-number</i>	N/A
3. Configure the load-sharing criteria for the aggregation group.	link-aggregation load-sharing mode { { destination-ip destination-mac source-ip source-mac } * }	The default load-sharing criteria are the same as the global load-sharing criteria.

In Layer 2 aggregate interface view, the switch supports the following load-sharing criteria and combinations:

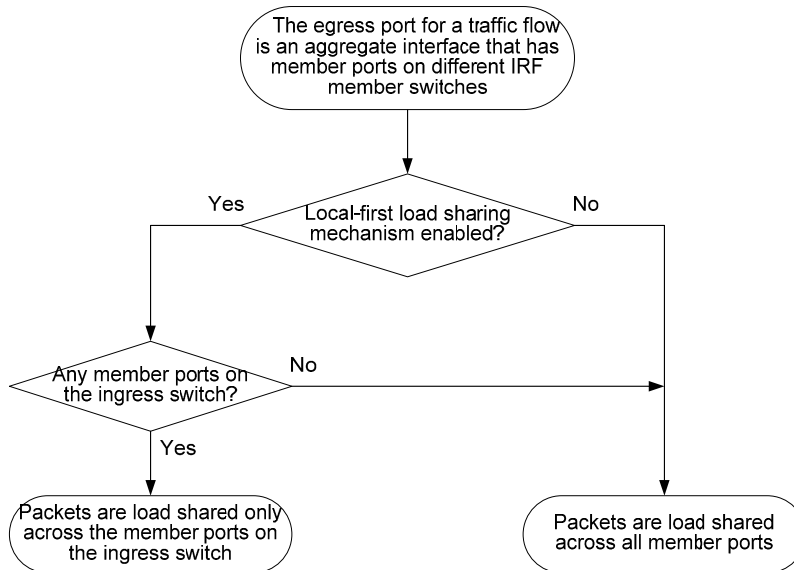
- Automatic load-sharing criteria determined based on the packet type
- Source IP address
- Destination IP address
- Source MAC address
- Destination MAC address
- Destination IP address and source IP address
- Destination MAC address and source MAC address

Enabling local-first load sharing for link aggregation

Use the local-first load sharing mechanism in a cross-switch link aggregation scenario to distribute traffic preferentially across member ports on the ingress switch rather than all member ports.

When you aggregate ports on different member switches in an IRF fabric, you can use local-first load sharing to reduce traffic on IRF links, as shown in [Figure 11](#). For more information about IRF, see *IRF Configuration Guide*.

Figure 11 Local-first link-aggregation load sharing



To enable local-first load sharing for link aggregation:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable local-first load-sharing for link aggregation.	link-aggregation load-sharing mode local-first	Optional Enabled by default.

Enabling link-aggregation traffic redirection

The link-aggregation traffic redirection function can redirect traffic between IRF member switches for a cross-device link aggregation group. Link-aggregation traffic redirection prevents traffic interruption when you reboot IRF member switch that contains link aggregation member ports. For more information about IRF, see *IRF Configuration Guide*.

Link-aggregation traffic redirection applies only to dynamic link aggregation groups and only to known unicast packets.

After link-aggregation traffic redirection is enabled, do not add an Ethernet interface configured with physical state change suppression to an aggregation group. Otherwise, Selected ports in the aggregation group might work improperly. For more information about physical state change suppression, see [Configuring Ethernet interfaces](#).

To enable link-aggregation traffic redirection:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable link-aggregation traffic redirection.	link-aggregation lacp traffic-redirect-notification enable	Optional Disabled by default.

△ CAUTION:

- To prevent traffic interruption, enable link-aggregation traffic redirection on devices at both ends of the aggregate link.
- To prevent packet loss that might occur at a reboot, disable both MSTP and link-aggregation traffic redirection.

Displaying and maintaining Ethernet link aggregation

Task	Command	Remarks
Display information about an aggregate interface or multiple aggregate interfaces.	display interface [bridge-aggregation route-aggregation] [brief [down]] [{ begin exclude include } <i>regular-expression</i>] display interface { bridge-aggregation route-aggregation } <i>interface-number</i> [brief] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the local system ID.	display lacp system-id [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the global or group-specific link-aggregation load-sharing criteria.	display link-aggregation load-sharing mode [interface [{ bridge-aggregation route-aggregation } <i>interface-number</i>]] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display detailed link aggregation information for link aggregation member ports.	display link-aggregation member-port [<i>interface-list</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display summary information about all aggregation groups.	display link-aggregation summary [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display detailed information about a specific or all aggregation groups.	display link-aggregation verbose [{ bridge-aggregation route-aggregation } [<i>interface-number</i>]] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear LACP statistics for a specific or all link aggregation member ports.	reset lacp statistics [interface <i>interface-list</i>]	Available in user view
Clear statistics for a specific or all aggregate interfaces.	reset counters interface [{ bridge-aggregation route-aggregation } [<i>interface-number</i>]]	Available in user view

Ethernet link aggregation configuration examples

In an aggregation group, only ports that have the same port attributes and class-two configurations (see "[Configuration classes](#)") as the reference port (see "[Reference port](#)") can operate as Selected ports. Make sure that all member ports have the same port attributes and class-two configurations as the reference port. The other settings only need to be configured on the aggregate interface, not on the member ports.

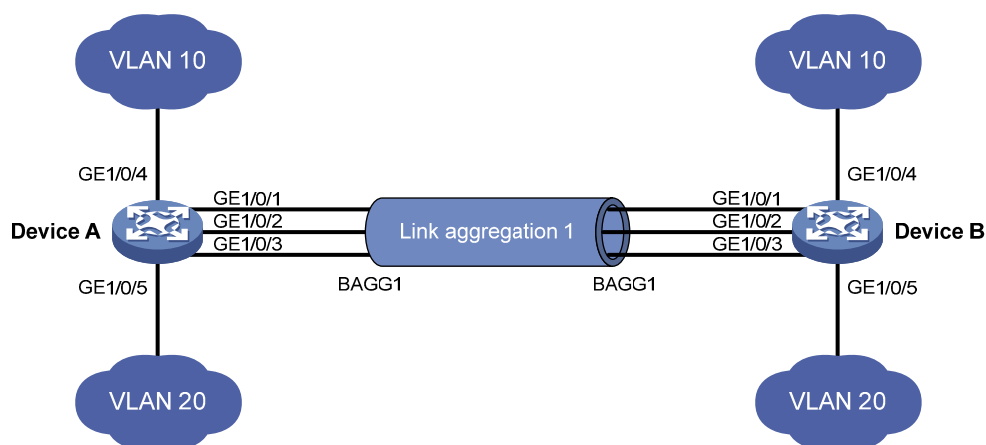
Layer 2 static aggregation configuration example

Network requirements

As shown in Figure 12:

- Device A and Device B are connected through their respective Layer 2 Ethernet interfaces GigabitEthernet 1/0/1 through GigabitEthernet 1/0/3.
- Configure a Layer 2 static aggregation group on both Device A and Device B. Enable VLAN 10 at one end of the aggregate link to communicate with VLAN 10 at the other end, and VLAN 20 at one end to communicate with VLAN 20 at the other end.
- Enable traffic to be load-shared across aggregation group member ports based on the source and destination MAC addresses.

Figure 12 Network diagram



Configuration procedure

1. Configure Device A:

Create VLAN 10, and assign port GigabitEthernet 1/0/4 to VLAN 10.

```
<DeviceA> system-view
[DeviceA] vlan 10
[DeviceA-vlan10] port gigabitethernet 1/0/4
[DeviceA-vlan10] quit
```

Create VLAN 20, and assign port GigabitEthernet 1/0/5 to VLAN 20.

```
[DeviceA] vlan 20
[DeviceA-vlan20] port gigabitethernet 1/0/5
[DeviceA-vlan20] quit
```

Create Layer 2 aggregate interface Bridge-Aggregation 1.

```
[DeviceA] interface bridge-aggregation 1
[DeviceA-Bridge-Aggregation1] quit
```

Assign ports GigabitEthernet 1/0/1 through GigabitEthernet 1/0/3 to link aggregation group 1.

```
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] port link-aggregation group 1
[DeviceA-GigabitEthernet1/0/1] quit
[DeviceA] interface gigabitethernet 1/0/2
```

```
[DeviceA-GigabitEthernet1/0/2] port link-aggregation group 1
[DeviceA-GigabitEthernet1/0/2] quit
[DeviceA] interface gigabitethernet 1/0/3
[DeviceA-GigabitEthernet1/0/3] port link-aggregation group 1
[DeviceA-GigabitEthernet1/0/3] quit
```

Configure Layer 2 aggregate interface Bridge-Aggregation 1 as a trunk port and assign it to VLANs 10 and 20.

```
[DeviceA] interface bridge-aggregation 1
[DeviceA-Bridge-Aggregation1] port link-type trunk
[DeviceA-Bridge-Aggregation1] port trunk permit vlan 10 20
Please wait... Done.
Configuring GigabitEthernet1/0/1... Done.
Configuring GigabitEthernet1/0/2... Done.
Configuring GigabitEthernet1/0/3... Done.
[DeviceA-Bridge-Aggregation1] quit
```

Configure Device A to use the source and destination MAC addresses of packets as the global link-aggregation load-sharing criteria.

```
[DeviceA] link-aggregation load-sharing mode source-mac destination-mac
```

2. Configure Device B.

Configure Device B using the same instructions that you used to configure Device A.

3. Verify the configurations:

Display summary information about all aggregation groups on Device A.

```
[DeviceA] display link-aggregation summary
```

```
Aggregation Interface Type:
BAGG -- Bridge-Aggregation, RAGG -- Route-Aggregation
Aggregation Mode: S -- Static, D -- Dynamic
Loadsharing Type: Shar -- Loadsharing, NonS -- Non-Loadsharing
Actor System ID: 0x8000, 000f-e2ff-0001
```

AGG	AGG	Partner ID	Select	Unselect	Share
Interface	Mode		Ports	Ports	Type
BAGG1	S	none	3	0	Shar

The output shows that link aggregation group 1 is a load-shared Layer 2 static aggregation group and it contains three Selected ports.

Display the global link-aggregation load-sharing criteria on Device A.

```
[DeviceA] display link-aggregation load-sharing mode
```

```
Link-Aggregation Load-Sharing Mode:
destination-mac address, source-mac address
```

The output shows that all link aggregation groups created on the device perform load sharing based on source and destination MAC addresses.

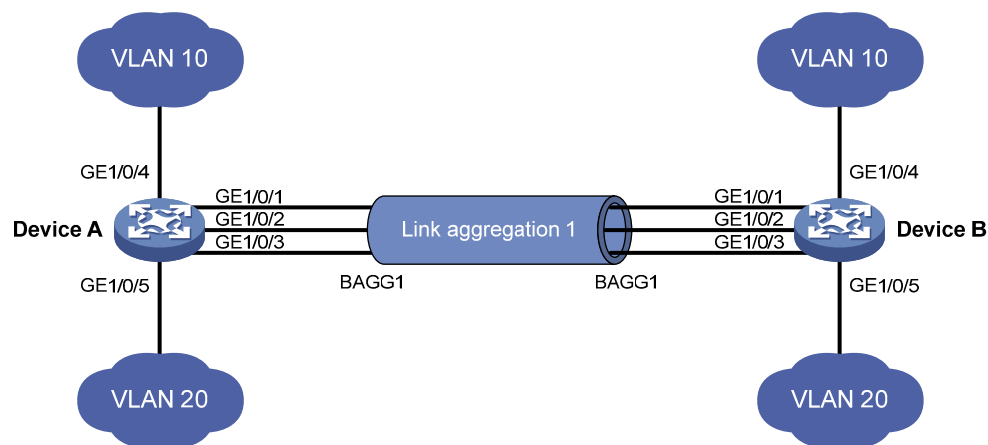
Layer 2 dynamic aggregation configuration example

Network requirements

As shown in Figure 13:

- Device A and Device B are connected through their respective Layer 2 Ethernet interfaces GigabitEthernet 1/0/1 through GigabitEthernet 1/0/3.
- Configure a Layer 2 dynamic aggregation group on both Device A and Device B, enable VLAN 10 at one end of the aggregate link to communicate with VLAN 10 at the other end, and VLAN 20 at one end to communicate with VLAN 20 at the other end.
- Enable traffic to be load-shared across aggregation group member ports based on source and destination MAC addresses.

Figure 13 Network diagram



Configuration procedure

1. Configure Device A:

Create VLAN 10, and assign the port GigabitEthernet 1/0/4 to VLAN 10.

```
<DeviceA> system-view
[DeviceA] vlan 10
[DeviceA-vlan10] port gigabitethernet 1/0/4
[DeviceA-vlan10] quit
```

Create VLAN 20, and assign the port GigabitEthernet 1/0/5 to VLAN 20.

```
[DeviceA] vlan 20
[DeviceA-vlan20] port gigabitethernet 1/0/5
[DeviceA-vlan20] quit
```

Create Layer 2 aggregate interface Bridge-Aggregation 1, and configure the link aggregation mode as dynamic.

```
[DeviceA] interface bridge-aggregation 1
[DeviceA-Bridge-Aggregation1] link-aggregation mode dynamic
```

Assign ports GigabitEthernet 1/0/1 through GigabitEthernet 1/0/3 to link aggregation group 1 one at a time.

```
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] port link-aggregation group 1
[DeviceA-GigabitEthernet1/0/1] quit
```



```
[DeviceA] interface gigabitethernet 1/0/2
[DeviceA-GigabitEthernet1/0/2] port link-aggregation group 1
[DeviceA-GigabitEthernet1/0/2] quit
[DeviceA] interface gigabitethernet 1/0/3
[DeviceA-GigabitEthernet1/0/3] port link-aggregation group 1
[DeviceA-GigabitEthernet1/0/3] quit
```

Configure Layer 2 aggregate interface Bridge-Aggregation 1 as a trunk port and assign it to VLANs 10 and 20.

```
[DeviceA] interface bridge-aggregation 1
[DeviceA-Bridge-Aggregation1] port link-type trunk
[DeviceA-Bridge-Aggregation1] port trunk permit vlan 10 20
Please wait... Done.
Configuring GigabitEthernet1/0/1... Done.
Configuring GigabitEthernet1/0/2... Done.
Configuring GigabitEthernet1/0/3... Done.
[DeviceA-Bridge-Aggregation1] quit
```

Configure the device to use the source and destination MAC addresses of packets as the global link-aggregation load-sharing criteria.

```
[DeviceA] link-aggregation load-sharing mode source-mac destination-mac
```

2. Configure Device B.

Configure Device B using the same instructions that you used to configure Device A.

3. Verify the configurations:

Display summary information about all aggregation groups on Device A.

```
[DeviceA] display link-aggregation summary
```

```
Aggregation Interface Type:
BAGG -- Bridge-Aggregation, RAGG -- Route-Aggregation
Aggregation Mode: S -- Static, D -- Dynamic
Loadsharing Type: Shar -- Loadsharing, NonS -- Non-Loadsharing
Actor System ID: 0x8000, 000f-e2ff-0001
```

AGG	AGG	Partner ID	Select	Unselect	Share
Interface	Mode		Ports	Ports	Type
BAGG1	D	0x8000, 000f-e2ff-0002	3	0	Shar

The output shows that link aggregation group 1 is a load-shared Layer 2 dynamic aggregation group and it contains three Selected ports.

Display the global link-aggregation load-sharing criteria on Device A.

```
[DeviceA] display link-aggregation load-sharing mode
```

```
Link-Aggregation Load-Sharing Mode:
destination-mac address, source-mac address
```

The output shows that all link aggregation groups created on the device perform load sharing based on source and destination MAC addresses.

Layer 3 static aggregation configuration example

Network requirements

As shown in Figure 14:

- Device A and Device B are connected by their Layer 3 Ethernet interfaces GigabitEthernet 1/0/1 through GigabitEthernet 1/0/3.
- Configure a Layer 3 static aggregation group on both Device A and Device B. Configure IP addresses and subnet masks for the corresponding Layer 3 aggregate interfaces.
- Enable traffic to be load-shared across aggregation group member ports based on source and destination IP addresses.

Figure 14 Network diagram



Configuration procedure

1. Configure Device A:

Create Layer 3 aggregate interface Route-Aggregation 1, and configure an IP address and subnet mask for the aggregate interface.

```
<DeviceA> system-view
[DeviceA] interface route-aggregation 1
[DeviceA-Route-Aggregation1] ip address 192.168.1.1 24
[DeviceA-Route-Aggregation1] quit
```

Assign Layer 3 Ethernet interfaces GigabitEthernet 1/0/1 through GigabitEthernet 1/0/3 to aggregation group 1.

```
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] port link-aggregation group 1
[DeviceA-GigabitEthernet1/0/1] quit
[DeviceA] interface gigabitethernet 1/0/2
[DeviceA-GigabitEthernet1/0/2] port link-aggregation group 1
[DeviceA-GigabitEthernet1/0/2] quit
[DeviceA] interface gigabitethernet 1/0/3
[DeviceA-GigabitEthernet1/0/3] port link-aggregation group 1
[DeviceA-GigabitEthernet1/0/3] quit
```

Configure the global link-aggregation load-sharing criteria as the source and destination IP addresses of packets.

```
[DeviceA] link-aggregation load-sharing mode source-ip destination-ip
```

2. Configure Device B.

Configure Device B using the same instructions that you used to configure Device A.

3. Verify the configurations:

Display summary information about all aggregation groups on Device A.

```
[DeviceA] display link-aggregation summary
```

Aggregation Interface Type:

```

BAGG -- Bridge-Aggregation, RAGG -- Route-Aggregation
Aggregation Mode: S -- Static, D -- Dynamic
Loadsharing Type: Shar -- Loadsharing, NonS -- Non-Loadsharing
Actor System ID: 0x8000, 000f-e2ff-0001

```

AGG	AGG	Partner ID	Select	Unselect	Share
Interface	Mode		Ports	Ports	Type
RAGG1	S	none	3	0	Shar

The output shows that link aggregation group 1 is a load-sharing-capable Layer 3 static aggregation group that contains three Selected ports.

```
# Display the global link-aggregation load-sharing criteria on Device A.
```

```
[DeviceA] display link-aggregation load-sharing mode
```

```
Link-Aggregation Load-Sharing Mode:
```

```
destination-ip address, source-ip address
```

The output shows that the global link-aggregation load-sharing criteria are the source and destination IP addresses of packets.

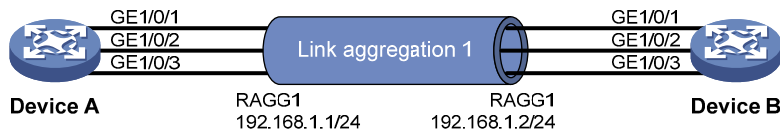
Layer 3 dynamic aggregation configuration example

Network requirements

As shown in [Figure 15](#):

- Device A and Device B are connected by their Layer 3 Ethernet interfaces GigabitEthernet 1/0/1 through GigabitEthernet 1/0/3.
- Configure a Layer 3 dynamic aggregation group on both Device A and Device B and configure IP addresses and subnet masks for the corresponding Layer 3 aggregate interfaces.
- Enable traffic to be load-shared across aggregation group member ports based on source and destination IP addresses.

Figure 15 Network diagram



Configuration procedure

1. Configure Device A:

```
# Create Layer 3 aggregate interface Route-Aggregation 1, configure the link aggregation mode as dynamic, and configure an IP address and subnet mask for the aggregate interface.
```

```

<DeviceA> system-view
[DeviceA] interface route-aggregation 1
[DeviceA-Route-Aggregation1] link-aggregation mode dynamic
[DeviceA-Route-Aggregation1] ip address 192.168.1.1 24
[DeviceA-Route-Aggregation1] quit

```

Assign Layer 3 Ethernet interfaces GigabitEthernet 1/0/1 through GigabitEthernet 1/0/3 to aggregation group 1.

```
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] port link-aggregation group 1
[DeviceA-GigabitEthernet1/0/1] quit
[DeviceA] interface gigabitethernet 1/0/2
[DeviceA-GigabitEthernet1/0/2] port link-aggregation group 1
[DeviceA-GigabitEthernet1/0/2] quit
[DeviceA] interface gigabitethernet 1/0/3
[DeviceA-GigabitEthernet1/0/3] port link-aggregation group 1
[DeviceA-GigabitEthernet1/0/3] quit
```

Configure Device A to use the source and destination IP addresses of packets as the global link-aggregation load-sharing criteria.

```
[DeviceA] link-aggregation load-sharing mode source-ip destination-ip
```

2. Configure Device B.

Configure Device B using the same instructions that you used to configure Device A.

3. Verify the configurations:

Display summary information about all aggregation groups on Device A.

```
[DeviceA] display link-aggregation summary
```

Aggregation Interface Type:

BAGG -- Bridge-Aggregation, RAGG -- Route-Aggregation

Aggregation Mode: S -- Static, D -- Dynamic

Loadsharing Type: Shar -- Loadsharing, NonS -- Non-Loadsharing

Actor System ID: 0x8000, 000f-e2ff-0001

AGG	AGG	Partner ID	Select	Unselect	Share
Interface	Mode		Ports	Ports	Type

RAGG1	D	0x8000, 000f-e2ff-0002	3	0	Shar

The output shows that link aggregation group 1 is a load-shared Layer 3 dynamic aggregation group and it contains three Selected ports.

Display the global link-aggregation load-sharing criteria on Device A.

```
[DeviceA] display link-aggregation load-sharing mode
```

Link-Aggregation Load-Sharing Mode:

```
destination-ip address, source-ip address
```

The output shows that the global link-aggregation load-sharing criteria are the source and destination IP addresses of packets.

Configuring port isolation

Port isolation enables isolating Layer 2 traffic for data privacy and security without using VLANs. You can also use this feature to isolate the hosts in a VLAN from one another.

To use the feature, assign ports to a port isolation group. Ports in an isolation group are called "isolated ports." One isolated port cannot forward Layer 2 traffic to any other isolated port on the same switch, even if they are in the same VLAN. An isolated port can communicate with any port outside the isolation group if they are in the same VLAN.

The switch series supports only one isolation group "isolation group 1." The isolation group is automatically created and cannot be deleted. There is no limit on the number of member ports.

Assigning a port to the isolation group

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none">Enter Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i>Enter Layer 2 aggregate interface view: interface bridge-aggregation <i>interface-number</i>Enter port group view: port-group manual <i>port-group-name</i>	<p>Use one of the commands.</p> <ul style="list-style-type: none">In Ethernet interface view, the subsequent configurations apply to the current port.In Layer 2 aggregate interface view, the subsequent configurations apply to the Layer 2 aggregate interface and all its member ports.In port group view, the subsequent configurations apply to all ports in the port group.
3. Assign the port or ports to the isolation group as an isolated port or ports.	port-isolate enable	No ports are added to the isolation group by default.

Displaying and maintaining the isolation group

Task	Command	Remarks
Display isolation group information.	display port-isolate group [{ begin exclude include } <i>regular-expression</i>]	Available in any view

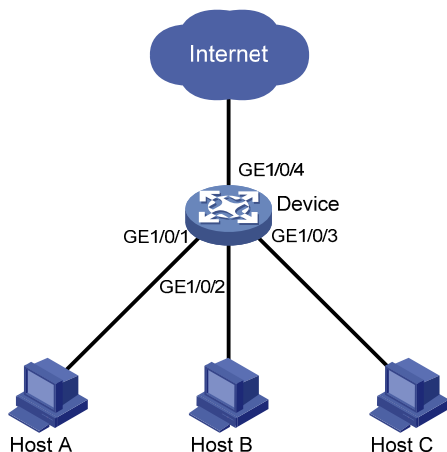
Port isolation configuration example

Network requirements

As shown in Figure 16, Host A, Host B, and Host C are connected to GigabitEthernet 1/0/1, GigabitEthernet 1/0/2, and GigabitEthernet 1/0/3 of Device, and Device is connected to the Internet through GigabitEthernet 1/0/4. All these ports are in the same VLAN.

Configure Device to provide Internet access for all the hosts and isolate them from one another.

Figure 16 Networking diagram



Configuration procedure

Add ports GigabitEthernet 1/0/1, GigabitEthernet 1/0/2, and GigabitEthernet 1/0/3 to the isolation group.

```
<Device> system-view
[Device] interface gigabitethernet 1/0/1
[Device-GigabitEthernet1/0/1] port-isolate enable
[Device-GigabitEthernet1/0/1] quit
[Device] interface gigabitethernet 1/0/2
[Device-GigabitEthernet1/0/2] port-isolate enable
[Device-GigabitEthernet1/0/2] quit
[Device] interface gigabitethernet 1/0/3
[Device-GigabitEthernet1/0/3] port-isolate enable
```

Display information about the isolation group.

```
<Device> display port-isolate group
Port-isolate group information:
Uplink port support: NO
Group ID: 1
Group members:
    GigabitEthernet 1/0/1    GigabitEthernet 1/0/2    GigabitEthernet 1/0/3
```

Configuring spanning tree protocols

As a Layer 2 management protocol, the Spanning Tree Protocol (STP) eliminates Layer 2 loops by selectively blocking redundant links in a network, putting them in a standby state, which still also allows for link redundancy.

The recent versions of STP include the Rapid Spanning Tree Protocol (RSTP), Per VLAN Spanning Tree (PVST), and the Multiple Spanning Tree Protocol (MSTP).

STP

STP was developed based on the 802.1d standard of IEEE to eliminate loops at the data link layer in a local area network (LAN). Networks often have redundant links as backups in case of failures, but loops are a very serious problem. Devices that run STP detect loops in the network by exchanging information with one another, and eliminate loops by selectively blocking certain ports to prune the loop structure into a loop-free tree structure. This avoids proliferation and infinite cycling of packets that would occur in a loop network, and prevents received duplicate packets from decreasing the performance of network devices.

In the narrow sense, STP refers to IEEE 802.1d STP. In the broad sense, STP refers to the IEEE 802.1d STP and various enhanced spanning tree protocols derived from that protocol.

STP protocol packets

STP uses bridge protocol data units (BPDUs), also known as configuration messages, as its protocol packets.

STP-enabled network devices exchange BPDUs to establish a spanning tree. BPDUs contain sufficient information for the network devices to complete spanning tree calculation.

STP uses the following types of BPDUs:

- **Configuration BPDUs**—Used by network devices to calculate a spanning tree and maintain the spanning tree topology
- **Topology change notification (TCN) BPDUs**—Notify network devices of the network topology changes

Configuration BPDUs contain sufficient information for the network devices to complete spanning tree calculation. Important fields in a configuration BPDU include the following:

- **Root bridge ID**—Consisting of the priority and MAC address of the root bridge.
- **Root path cost**—Cost of the path to the root bridge denoted by the root identifier from the transmitting bridge.
- **Designated bridge ID**—Consisting of the priority and MAC address of the designated bridge.
- **Designated port ID**—Consisting of the priority and global port number of the designated port.
- **Message age**—Age of the configuration BPDU while it propagates in the network.
- **Max age**—Maximum age of the configuration BPDU stored on the switch.
- **Hello time**—Configuration BPDU transmission interval.
- **Forward delay**—Delay that STP bridges use to transition port state.

Basic concepts in STP

Root bridge

A tree network must have a root bridge. The entire network contains only one root bridge. The root bridge is not permanent, but can change with changes of the network topology.

Upon initialization of a network, each device generates and periodically sends configuration BPDUs with itself as the root bridge. After network convergence, only the root bridge generates and periodically sends configuration BPDUs, and the other devices forward the BPDUs.

Root port

On a non-root bridge, the port nearest to the root bridge is the root port. The root port communicates with the root bridge. Each non-root bridge has only one root port. The root bridge has no root port.

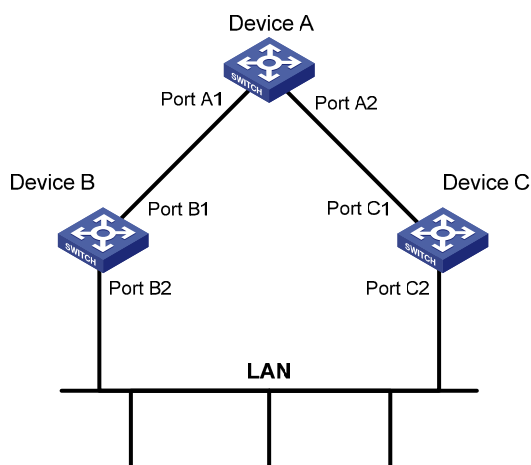
Designated bridge and designated port

Table 8 Description of designated bridges and designated ports

Classification	Designated bridge	Designated port
For a device	Device directly connected with the local device and responsible for forwarding BPDUs to the local device	Port through which the designated bridge forwards BPDUs to this device
For a LAN	Device responsible for forwarding BPDUs to this LAN segment	Port through which the designated bridge forwards BPDUs to this LAN segment

As shown in Figure 17, Device B and Device C are directly connected to a LAN. If Device A forwards BPDUs to Device B through port A1, the designated bridge for Device B is Device A, and the designated port of Device B is port A1 on Device A. If Device B forwards BPDUs to the LAN, the designated bridge for the LAN is Device B, and the designated port for the LAN is port B2 on Device B.

Figure 17 Designated bridges and designated ports



Path cost

Path cost is a reference value used for link selection in STP. STP calculates path costs to select the most robust links and block redundant links that are less robust, to prune the network into a loop-free tree.

Calculation process of the STP algorithm

The spanning tree calculation process described in the following sections is a simplified process for example only.

The STP algorithm uses the following calculation process:

1. Initial state

Upon initialization of a device, each port generates a BPDU with the port as the designated port, the device as the root bridge, 0 as the root path cost, and the device ID as the designated bridge ID.

2. Root bridge selection

Initially, each STP-enabled device on the network assumes itself to be the root bridge, with its own device ID as the root bridge ID. By exchanging configuration BPDUs, the devices compare their root bridge IDs to elect the device with the smallest root bridge ID as the root bridge.

3. Non-root bridge: Selection of root port and designated ports

Table 9 Selection of the root port and designated ports

Step	Description
1	A non-root-bridge device regards the port on which it received the optimum configuration BPDU as the root port. Table 10 describes how the optimum configuration BPDU is selected.
2	Based on the configuration BPDU and the path cost of the root port, the device calculates a designated port configuration BPDU for each of the other ports. <ul style="list-style-type: none">• The root bridge ID is replaced with that of the configuration BPDU of the root port.• The root path cost is replaced with that of the configuration BPDU of the root port plus the path cost of the root port.• The designated bridge ID is replaced with the ID of this device.• The designated port ID is replaced with the ID of this port.
3	The device compares the calculated configuration BPDU with the configuration BPDU on the port whose port role will be defined, and acts depending on the result of the comparison. <ul style="list-style-type: none">• If the calculated configuration BPDU is superior, the device considers this port as the designated port, replaces the configuration BPDU on the port with the calculated configuration BPDU, and periodically sends the calculated configuration BPDU.• If the configuration BPDU on the port is superior, the device blocks this port without updating its configuration BPDU. The blocked port can receive BPDUs, but cannot send BPDUs or forward data traffic.

NOTE:

When the network topology is stable, only the root port and designated ports forward user traffic, while other ports are all in the blocked state to receive BPDUs but not forward BPDUs or user traffic.

Table 10 Selection of the optimum configuration BPDU

Step	Actions
1	<p>Upon receiving a configuration BPDU on a port, the device compares the priority of the received configuration BPDU with that of the configuration BPDU generated by the port, and:</p> <ul style="list-style-type: none"> • If the former priority is lower, the device discards the received configuration BPDU and keeps the configuration BPDU the port generated. • If the former priority is higher, the device replaces the content of the configuration BPDU generated by the port with the content of the received configuration BPDU.
2	The device compares the configuration BPDUs of all the ports and chooses the optimum configuration BPDU.

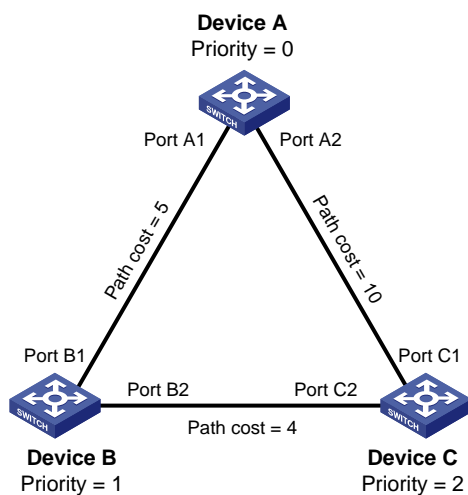
The following are the principles of configuration BPDU comparison:

- The configuration BPDU with the lowest root bridge ID has the highest priority.
- If configuration BPDUs have the same root bridge ID, their root path costs are compared. For example, the root path cost in a configuration BPDU plus the path cost of a receiving port is S. The configuration BPDU with the smallest S value has the highest priority.
- If all configuration BPDUs have the same ports value, their designated bridge IDs, designated port IDs, and the IDs of the receiving ports are compared in sequence. The configuration BPDU that contains the smallest ID wins.

A tree-shape topology forms when the root bridge, root ports, and designated ports are selected.

Figure 18 describes with an example how the STP algorithm works. This example shows a simplified spanning tree calculation process.

Figure 18 The STP algorithm



As shown in Figure 18, the priority values of Device A, Device B, and Device C are 0, 1, and 2, and the path costs of links among the three devices are 5, 10, and 4, respectively.

4. Initial state of each device

Table 11 Initial state of each device

Device	Port name	Configuration BPDU on the port
Device A	Port A1	{0, 0, 0, Port A1}
	Port A2	{0, 0, 0, Port A2}
Device B	Port B1	{1, 0, 1, Port B1}
	Port B2	{1, 0, 1, Port B2}
Device C	Port C1	{2, 0, 2, Port C1}
	Port C2	{2, 0, 2, Port C2}

NOTE:

In Table 11, each configuration BPDU contains the following fields: root bridge ID, root path cost, designated bridge ID, and designated port ID.

5. Comparison process and result on each device

Table 12 Comparison process and result on each device

Device	Comparison process	Configuration BPDU on ports after comparison
Device A	<ul style="list-style-type: none"> Port A1 receives the configuration BPDU of Port B1 {1, 0, 1, Port B1}, finds that its existing configuration BPDU {0, 0, 0, Port A1} is superior to the received configuration BPDU, and discards the received one. Port A2 receives the configuration BPDU of Port C1 {2, 0, 2, Port C1}, finds that its existing configuration BPDU {0, 0, 0, Port A2} is superior to the received configuration BPDU, and discards the received one. Device A finds that it is both the root bridge and designated bridge in the configuration BPDUs of all its ports, and considers itself as the root bridge. It does not change the configuration BPDU of any port and starts to periodically send configuration BPDUs. 	<ul style="list-style-type: none"> Port A1: {0, 0, 0, Port A1} Port A2: {0, 0, 0, Port A2}
Device B	<ul style="list-style-type: none"> Port B1 receives the configuration BPDU of Port A1 {0, 0, 0, Port A1}, finds that the received configuration BPDU is superior to its existing configuration BPDU {1, 0, 1, Port B1}, and updates its configuration BPDU. Port B2 receives the configuration BPDU of Port C2 {2, 0, 2, Port C2}, finds that its existing configuration BPDU {1, 0, 1, Port B2} is superior to the received configuration BPDU, and discards the received one. 	<ul style="list-style-type: none"> Port B1: {0, 0, 0, Port A1} Port B2: {1, 0, 1, Port B2}

Device	Comparison process	Configuration BPDU on ports after comparison
	<ul style="list-style-type: none"> Device B compares the configuration BPDUs of all its ports, decides that the configuration BPDU of Port B1 is the optimum, and selects Port B1 as the root port with the configuration BPDU unchanged. Based on the configuration BPDU and path cost of the root port, Device B calculates a designated port configuration BPDU for Port B2 {0, 5, 1, Port B2}, and compares it with the existing configuration BPDU of Port B2 {1, 0, 1, Port B2}. Device B finds that the calculated one is superior, decides that Port B2 is the designated port, replaces the configuration BPDU on Port B2 with the calculated one, and periodically sends the calculated configuration BPDU. 	<ul style="list-style-type: none"> Root port (Port B1): {0, 0, 0, Port A1} Designated port (Port B2): {0, 5, 1, Port B2}
	<ul style="list-style-type: none"> Port C1 receives the configuration BPDU of Port A2 {0, 0, 0, Port A2}, finds that the received configuration BPDU is superior to its existing configuration BPDU {2, 0, 2, Port C1}, and updates its configuration BPDU. Port C2 receives the original configuration BPDU of Port B2 {1, 0, 1, Port B2}, finds that the received configuration BPDU is superior to the existing configuration BPDU {2, 0, 2, Port C2}, and updates its configuration BPDU. 	<ul style="list-style-type: none"> Port C1: {0, 0, 0, Port A2} Port C2: {1, 0, 1, Port B2}
Device C	<ul style="list-style-type: none"> Device C compares the configuration BPDUs of all its ports, decides that the configuration BPDU of Port C1 is the optimum, and selects Port C1 as the root port with the configuration BPDU unchanged. Based on the configuration BPDU and path cost of the root port, Device C calculates the configuration BPDU of Port C2 {0, 10, 2, Port C2}, and compares it with the existing configuration BPDU of Port C2 {1, 0, 1, Port B2}. Device C finds that the calculated configuration BPDU is superior to the existing one, selects Port C2 as the designated port, and replaces the configuration BPDU of Port C2 with the calculated one. 	<ul style="list-style-type: none"> Root port (Port C1): {0, 0, 0, Port A2} Designated port (Port C2): {0, 10, 2, Port C2}
	<ul style="list-style-type: none"> Port C2 receives the updated configuration BPDU of Port B2 {0, 5, 1, Port B2}, finds that the received configuration BPDU is superior to its existing configuration BPDU {0, 10, 2, Port C2}, and updates its configuration BPDU. Port C1 receives a periodic configuration BPDU {0, 0, 0, Port A2} from Port A2, finds that it is the same as the existing configuration BPDU, and discards the received one. 	<ul style="list-style-type: none"> Port C1: {0, 0, 0, Port A2} Port C2: {0, 5, 1, Port B2}

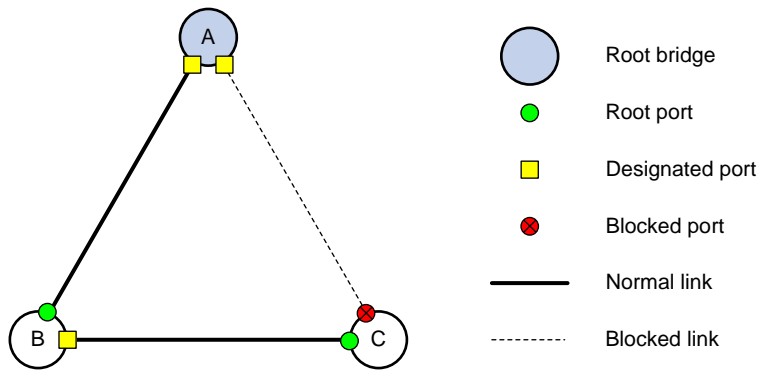
Device	Comparison process	Configuration BPDU on ports after comparison
	<ul style="list-style-type: none"> Device C finds that the root path cost of Port C1 (10) (root path cost of the received configuration BPDU (0) plus path cost of Port C1 (10)) is larger than that of Port C2 (9) (root path cost of the received configuration BPDU (5) plus path cost of Port C2 (4)), decides that the configuration BPDU of Port C2 is the optimum, and selects Port C2 as the root port with the configuration BPDU unchanged. Based on the configuration BPDU and path cost of the root port, Device C calculates a designated port configuration BPDU for Port C1 {0, 9, 2, Port C1} and compares it with the existing configuration BPDU of Port C1 {0, 0, 0, Port A2}. Device C finds that the existing configuration BPDU is superior to the calculated one and blocks Port C1 with the configuration BPDU unchanged. Then Port C1 does not forward data until a new event triggers a spanning tree calculation process, for example, the link between Device B and Device C is down. 	<ul style="list-style-type: none"> Blocked port (Port C1): {0, 0, 0, Port A2} Root port (Port C2): {0, 5, 1, Port B2}

NOTE:

In Table 12, each configuration BPDU contains the following fields: root bridge ID, root path cost, designated bridge ID, and designated port ID.

After the comparison processes described in Table 12, a spanning tree with Device A as the root bridge is established, and the topology is shown in Figure 19.

Figure 19 The final calculated spanning tree



The configuration BPDU forwarding mechanism of STP

The configuration BPDUs of STP are forwarded following these guidelines:

- Upon network initiation, every device regards itself as the root bridge, generates configuration BPDUs with itself as the root, and sends the configuration BPDUs at a regular hello interval.
- If the root port received a configuration BPDU and the received configuration BPDU is superior to the configuration BPDU of the port, the device increases the message age carried in the configuration BPDU following a certain rule and starts a timer to time the configuration BPDU while sending this configuration BPDU through the designated port.
- If the configuration BPDU received on a designated port has a lower priority than the configuration BPDU of the local port, the port immediately sends its own configuration BPDU in response.

- If a path becomes faulty, the root port on this path no longer receives new configuration BPDUs and the old configuration BPDUs will be discarded due to timeout. The device generates a configuration BPDU with itself as the root and sends the BPDUs and TCN BPDUs. This triggers a new spanning tree calculation process to establish a new path to restore the network connectivity.

However, the newly calculated configuration BPDU cannot be propagated throughout the network immediately, so the old root ports and designated ports that have not detected the topology change continue forwarding data along the old path. If the new root ports and designated ports begin to forward data as soon as they are elected, a temporary loop might occur.

STP timers

The most important timing parameters in STP calculation are forward delay, hello time, and max age.

- Forward delay

Forward delay is the delay time for port state transition.

A path failure can cause spanning tree re-calculation to adapt the spanning tree structure to the change. However, the resulting new configuration BPDU cannot propagate throughout the network immediately. If the newly elected root ports and designated ports start to forward data immediately, a temporary loop will likely occur.

For this reason, as a mechanism for state transition in STP, the newly elected root ports or designated ports require twice the forward delay time before they transit to the forwarding state to make sure that the new configuration BPDU has propagated throughout the network.

- Hello time

The device sends hello packets at the hello time interval to the neighboring devices to make sure that the paths are fault-free.

- Max age

The device uses the max age to determine whether a stored configuration BPDU has expired and discards it if the max age is exceeded.

RSTP

RSTP achieves rapid network convergence by allowing a newly elected root port or designated port to enter the forwarding state much faster under certain conditions than STP.

A newly elected RSTP root port rapidly enters the forwarding state if the old root port on the device has stopped forwarding data and the upstream designated port has started forwarding data.

A newly elected RSTP designated port rapidly enters the forwarding state if it is an edge port (a port that directly connects to a user terminal rather than to another network device or a shared LAN segment) or it connects to a point-to-point link (to another device). Edge ports directly enter the forwarding state. Connecting to a point-to-point link, a designated port enters the forwarding state immediately after the device receives a handshake response from the directly connected device.

PVST

PVST was introduced to improve link bandwidth usage in network environments where multiple virtual LANs (VLANs) exist. Unlike STP and RSTP whose bridges in a LAN must forward their VLAN packets in the same spanning tree, PVST allows each VLAN to build a separate spanning tree.

PVST uses the following BPDUs:

- **STP BPDUs**—Sent by access ports according to the VLAN status, or by trunk ports and hybrid ports according to the status of VLAN 1.
- **PVST BPDUs**—Sent by trunk port and hybrid ports according to the status of permitted VLANs except VLAN 1.

MSTP

STP, RSTP, and PVST limitations

STP does not support rapid state transition of ports. A newly elected port must wait twice the forward delay time before it transits to the forwarding state, even if it connects to a point-to-point link or is an edge port.

Although RSTP supports rapid network convergence, it has the same drawback as STP—All bridges within a LAN share the same spanning tree, so redundant links cannot be blocked based on VLAN, and the packets of all VLANs are forwarded along the same spanning tree.

The number of PVST BPDUs generated grows with that of permitted VLANs on trunk ports. When the status of a trunk port transitions, network devices might be overloaded to re-calculate a large number of spanning trees.

MSTP features

Developed based on IEEE 802.1s, MSTP overcomes the limitations of STP, RSTP, and PVST. In addition to supporting rapid network convergence, it provides a better load sharing mechanism for redundant links by allowing data flows of different VLANs to be forwarded along separate paths.

MSTP provides the following features:

- MSTP supports mapping VLANs to spanning tree instances by means of a VLAN-to-instance mapping table. MSTP can reduce communication overheads and resource usage by mapping multiple VLANs to one instance.
- MSTP divides a switched network into multiple regions, each of which contains multiple spanning trees that are independent of one another.
- MSTP prunes a loop network into a loop-free tree, which avoids proliferation and endless cycling of packets in a loop network. In addition, it supports load balancing of VLAN data by providing multiple redundant paths for data forwarding.

MSTP basic concepts

Figure 20 shows a switched network that comprises four MST regions, each MST region comprising four MSTP devices. Figure 21 shows the networking topology of MST region 3.

Figure 20 Basic concepts in MSTP

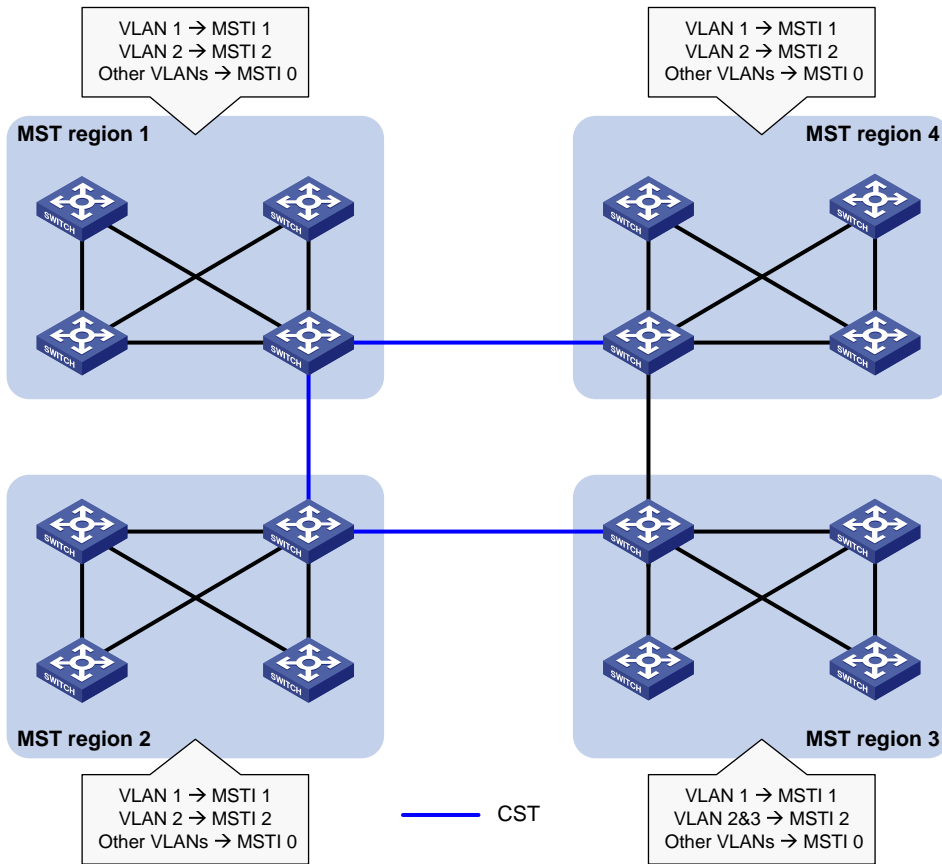
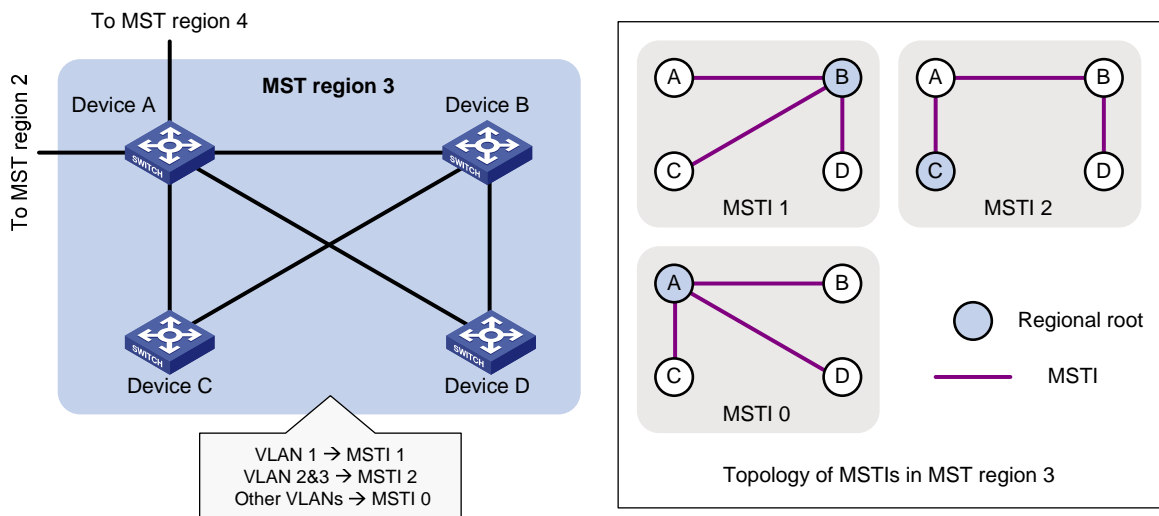


Figure 21 Network diagram and topology of MST region 3



MST region

A multiple spanning tree region (MST region) consists of multiple devices in a switched network and the network segments among them. All these devices have the following characteristics:

- A spanning tree protocol enabled
- Same region name

- Same VLAN-to-instance mapping configuration
- Same MSTP revision level
- Physically linked together

Multiple MST regions can exist in a switched network. You can assign multiple devices to the same MST region. In [Figure 20](#), the switched network comprises four MST regions, MST region 1 through MST region 4, and all devices in each MST region have the same MST region configuration.

MSTI

MSTP can generate multiple independent spanning trees in an MST region, and each spanning tree is mapped to the specific VLANs. Each spanning tree is referred to as a multiple spanning tree instance (MSTI).

In [Figure 21](#), MST region 3 comprises three MSTIs, MSTI 1, MSTI 2, and MSTI 0.

VLAN-to-instance mapping table

As an attribute of an MST region, the VLAN-to-instance mapping table describes the mapping relationships between VLANs and MSTIs.

In [Figure 21](#), the VLAN-to-instance mapping table of MST region 3 is: VLAN 1 to MSTI 1, VLAN 2 and VLAN 3 to MSTI 2, and other VLANs to MSTI 0. MSTP achieves load balancing by means of the VLAN-to-instance mapping table.

CST

The common spanning tree (CST) is a single spanning tree that connects all MST regions in a switched network. If you regard each MST region as a device, the CST is a spanning tree calculated by these devices through STP or RSTP.

The blue lines in [Figure 20](#) represent the CST.

IST

An internal spanning tree (IST) is a spanning tree that runs in an MST region. It is also called MSTI 0, a special MSTI to which all VLANs are mapped by default.

In [Figure 20](#), MSTI 0 is the IST in MST region 3.

CIST

The common and internal spanning tree (CIST) is a single spanning tree that connects all devices in a switched network. It consists of the ISTs in all MST regions and the CST.

In [Figure 20](#), the ISTs (MSTI 0) in all MST regions plus the inter-region CST constitute the CIST of the entire network.

Regional root

The root bridge of the IST or an MSTI within an MST region is the regional root of the IST or MSTI. Based on the topology, different spanning trees in an MST region might have different regional roots.

For example, in MST region 3 in [Figure 21](#), the regional root of MSTI 1 is Device B, the regional root of MSTI 2 is Device C, and the regional root of MSTI 0 (also known as the IST) is Device A.

Common root bridge

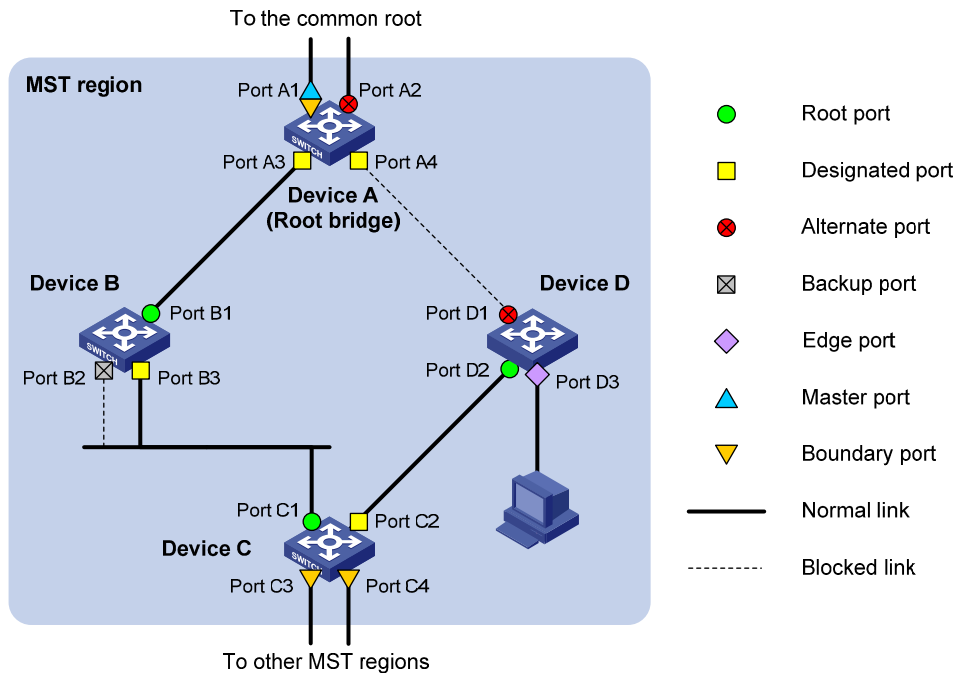
The common root bridge is the root bridge of the CIST.

In [Figure 20](#), for example, the common root bridge is a device in MST region 1.

Port roles

A port can play different roles in different MSTIs. As shown in [Figure 22](#), an MST region comprises Device A, Device B, Device C, and Device D. Port A1 and port A2 of Device A connect to the common root bridge. Port B2 and Port B3 of Device B form a loop. Port C3 and Port C4 of Device C connect to other MST regions. Port D3 of Device D directly connects to a host.

Figure 22 Port roles



MSTP calculation involves the following port roles:

- **Root port**—Forwards data for a non-root bridge to the root bridge. The root bridge does not have any root port.
- **Designated port**—Forwards data to the downstream network segment or device.
- **Alternate port**—The backup port for a root port or master port. When the root port or master port is blocked, the alternate port takes over.
- **Backup port**—The backup port of a designated port. When the designated port is invalid, the backup port becomes the new designated port. A loop occurs when two ports of the same spanning tree device are interconnected, so the device blocks one of the ports. The blocked port acts as the backup.
- **Edge port**—An edge port does not connect to any network device or network segment, but directly connects to a user host.
- **Master port**—A port on the shortest path from the local MST region to the common root bridge. The master port is not always located on the regional root. It is a root port on the IST or CIST and still a master port on the other MSTIs.
- **Boundary port**—Connects an MST region to another MST region or to an STP/RSTP-running device. In MSTP calculation, a boundary port's role on an MSTI is consistent with its role on the CIST. But that is not true with master ports. A master port on MSTIs is a root port on the CIST.

Port states

In MSTP, a port can be in one of the following states:

- **Forwarding**—The port receives and sends BPDUs, obtains MAC addresses, and forwards user traffic.
- **Learning**—The port receives and sends BPDUs, obtains MAC addresses, but does not forward user traffic. Learning is an intermediate port state.
- **Discarding**—The port receives and sends BPDUs, but does not obtain MAC addresses or forward user traffic.

When in different MSTIs, a port can be in different states. A port state is not exclusively associated with a port role. Table 13 lists the port states that each port role supports. (A check mark [✓] indicates that the port supports this state, while a dash [—] indicates that the port does not support this state.)

Table 13 Port states that different port roles support

Port state (below)	Port role (right) Root port/master port	Designated port	Alternate port	Backup port
Forwarding	✓	✓	—	—
Learning	✓	✓	—	—
Discarding	✓	✓	✓	✓

How MSTP works

MSTP divides an entire Layer 2 network into multiple MST regions, which are connected by a calculated CST. Inside an MST region, multiple spanning trees are calculated. Each spanning tree is an MSTI. Among these MSTIs, MSTI 0 is the IST. Like STP, MSTP uses configuration BPDUs to calculate spanning trees. An important difference is that an MSTP BPDU carries the MSTP configuration of the bridge from which the BPDU is sent.

CIST calculation

The calculation of a CIST tree is also the process of configuration BPDU comparison. During this process, the device with the highest priority is elected as the root bridge of the CIST. MSTP generates an IST within each MST region through calculation. At the same time, MSTP regards each MST region as a single device and generates a CST among these MST regions through calculation. The CST and ISTs constitute the CIST of the entire network.

MSTI calculation

Within an MST region, MSTP generates different MSTIs for different VLANs based on the VLAN-to-instance mappings. For each spanning tree, MSTP performs a separate calculation process similar to spanning tree calculation in STP. For more information, see "[Calculation process of the STP algorithm.](#)"

In MSTP, a VLAN packet is forwarded along the following paths:

- Within an MST region, the packet is forwarded along the corresponding MSTI.
- Between two MST regions, the packet is forwarded along the CST.

Implementation of MSTP on devices

MSTP is compatible with STP and RSTP. Devices that are running MSTP and that are used for spanning tree calculation can identify STP and RSTP protocol packets.

In addition to basic MSTP functions, the following functions are provided for ease of management:

- Root bridge hold
- Root bridge backup
- Root guard
- BPDU guard
- Loop guard
- TC-BPDU guard
- BPDU drop.

Protocols and standards

The spanning tree protocols are documented in the following standards:

- IEEE 802.1d, *Media Access Control (MAC) Bridges*
- IEEE 802.1w, *Part 3: Media Access Control (MAC) Bridges—Amendment 2: Rapid Reconfiguration*
- IEEE 802.1s, *Virtual Bridged Local Area Networks—Amendment 3: Multiple Spanning Trees*

Spanning tree configuration task list

Before configuring a spanning tree, you must determine the spanning tree protocol to be used (STP, RSTP, PVST, or MSTP) and plan the device roles (the root bridge or leaf node).

Configuration restrictions and guidelines

- If GVRP and a spanning tree protocol are enabled on a device at the same time, GVRP packets are forwarded along the CIST. To advertise a certain VLAN within the network through GVRP, be sure that this VLAN is mapped to the CIST when you configure the VLAN-to-instance mapping table. For more information about GVRP, see "[Configuring GVRP](#)."
- The spanning tree configurations are mutually exclusive with any of the following functions on a port: service loopback, RRPP, Smart Link, and BPDU tunneling for STP.
- The spanning tree configurations made in system view take effect globally. Configurations made in Layer 2 Ethernet interface view take effect on the current interface only. Configurations made in port group view take effect on all member ports in the port group. Configurations made in Layer 2 aggregate interface view take effect only on the aggregate interface. Configurations made on an aggregation member port can take effect only after the port is removed from the aggregation group.
- After you enable a spanning tree protocol on a Layer 2 aggregate interface, the system performs spanning tree calculation on the Layer 2 aggregate interface but not on the aggregation member ports. The spanning tree protocol enable state and forwarding state of each selected member port is consistent with those of the corresponding Layer 2 aggregate interface.
- Though the member ports of an aggregation group do not participate in spanning tree calculation, the ports still reserve their spanning tree configurations for participating in spanning tree calculation after leaving the aggregation group.

STP configuration task list

Task	Remarks
Configuring the root bridge	Setting the spanning tree mode Required Configure the device to operate in STP-compatible mode.
	Configuring the root bridge or a secondary root bridge Optional
	Configuring the device priority Optional
	Configuring the network diameter of a switched network Optional
	Configuring spanning tree timers Optional
	Configuring the timeout factor Optional
	Configuring the maximum port rate Optional
	Configuring the mode a port uses to recognize/send MSTP packets Optional
	Enabling outputting port state transition information Optional
	Enabling the spanning tree feature Required
Configuring the leaf nodes	Setting the spanning tree mode Required Configure the device to operate in STP-compatible mode.
	Configuring the device priority Optional
	Configuring the timeout factor Optional
	Configuring the maximum port rate Optional
	Configuring path costs of ports Optional
	Configuring the port priority Optional
	Configuring the mode a port uses to recognize/send MSTP packets Optional
	Enabling outputting port state transition information Optional
Enabling the spanning tree feature Required	
Configuring TC snooping Optional	
Configuring protection functions Optional	

RSTP configuration task list

Task	Remarks
Configuring the root bridge	Setting the spanning tree mode Required Configure the device to operate in RSTP mode.
	Configuring the root bridge or a secondary root bridge Optional
	Configuring the device priority Optional
	Configuring the network diameter of a switched network Optional

Task	Remarks
Configuring spanning tree timers	Optional
Configuring the timeout factor	Optional
Configuring the maximum port rate	Optional
Configuring edge ports	Optional
Configuring the port link type	Optional
Configuring the mode a port uses to recognize/send MSTP packets	Optional
Enabling outputting port state transition information	Optional
Enabling the spanning tree feature	Required
Setting the spanning tree mode	Required Configure the device to operate in RSTP mode.
Configuring the device priority	Optional
Configuring the timeout factor	Optional
Configuring the maximum port rate	Optional
Configuring edge ports	Optional
Configuring path costs of ports	Optional
Configuring the port priority	Optional
Configuring the port link type	Optional
Configuring the mode a port uses to recognize/send MSTP packets	Optional
Enabling outputting port state transition information	Optional
Enabling the spanning tree feature	Required
Performing mCheck	Optional
Configuring TC snooping	Optional
Configuring protection functions	Optional

Configuring the leaf nodes

PVST configuration task list

Task	Remarks
Setting the spanning tree mode	Required Configure the device to operate in PVST mode.
Configuring the root bridge or a secondary root bridge	Optional
Configuring the device priority	Optional
Configuring the network diameter of a switched network	Optional

Task	Remarks
Configuring spanning tree timers	Optional
Configuring the timeout factor	Optional
Configuring the maximum port rate	Optional
Configuring edge ports	Optional
Configuring the port link type	Optional
Enabling outputting port state transition information	Optional
Enabling the spanning tree feature	Required
Setting the spanning tree mode	Required Configure the device to operate in PVST mode.
Configuring the device priority	Optional
Configuring the timeout factor	Optional
Configuring the maximum port rate	Optional
Configuring edge ports	Optional
Configuring path costs of ports	Optional
Configuring the port priority	Optional
Configuring the port link type	Optional
Enabling outputting port state transition information	Optional
Enabling the spanning tree feature	Required
Performing mCheck	Optional
Configuring protection functions	Optional

Configuring the leaf nodes

MSTP configuration task list

Task	Remarks
Setting the spanning tree mode	Optional By default, the device operates in MSTP mode.
Configuring an MST region	Required
Configuring the root bridge or a secondary root bridge	Optional
Configuring the device priority	Optional
Configuring the maximum hops of an MST region	Optional
Configuring the network diameter of a switched network	Optional
Configuring spanning tree timers	Optional
Configuring the timeout factor	Optional

Configuring the root bridge

Task	Remarks	
Configuring the maximum port rate	Optional	
Configuring edge ports	Optional	
Configuring the port link type	Optional	
Configuring the mode a port uses to recognize/send MSTP packets	Optional	
Enabling outputting port state transition information	Optional	
Enabling the spanning tree feature	Required	
Setting the spanning tree mode	Optional By default, the device operates in MSTP mode.	
Configuring an MST region	Required	
Configuring the device priority	Optional	
Configuring the timeout factor	Optional	
Configuring the leaf nodes	Configuring the maximum port rate	Optional
	Configuring edge ports	Optional
	Configuring path costs of ports	Optional
	Configuring the port priority	Optional
	Configuring the port link type	Optional
	Configuring the mode a port uses to recognize/send MSTP packets	Optional
	Enabling outputting port state transition information	Optional
	Enabling the spanning tree feature	Required
	Performing mCheck	Optional
	Configuring Digest Snooping	Optional
Configuring No Agreement Check	Optional	
Configuring protection functions	Optional	

Setting the spanning tree mode

The spanning tree modes include:

- **STP-compatible mode**—The device sends STP BPDUs through all ports.
- **RSTP mode**—The device sends RSTP BPDUs through all ports, and ports that connect to STP devices automatically transitions to the STP-compatible mode.
- **MSTP mode**—The device sends MSTP BPDUs through all ports, and ports that connect to STP devices automatically transitions to the STP-compatible mode.
- **PVST mode**—The device sends PVST BPDUs through all ports and maintains a spanning tree for each VLAN. The number of VLANs that PVST can maintain instances for depends on the switch model. Suppose the number is n , which is 128 on the 5500 EI Switch Series and 32 on the 5500 SI Switch Series. When you configure PVST on devices of different models in a network, to avoid network failures, make sure that the number of VLANs for which PVST maintains instances does not exceed the lowest n .

The MSTP mode is compatible with the RSTP mode, and the RSTP mode is compatible with the STP mode. The PVST mode's compatibility with the other spanning tree mode varies by port type:

- On an access port, the PVST mode is compatible with any other spanning tree mode in any VLAN.
- On a trunk or hybrid port, the PVST mode is compatible with any other spanning tree mode in only VLAN 1.

Whether you need to specify the MSTI or VLAN for the spanning tree configuration varies with the spanning tree modes.

- In STP-compatible or RSTP mode, do not specify any MSTI or VLAN. Otherwise, the spanning tree configuration is ineffective.
- In MSTP mode, if you specify an MSTI, the spanning tree configuration is effective for the specified MSTI. If you specify a VLAN list, the spanning tree configuration is ineffective. If you do not specify any MSTI or VLAN, the spanning tree configuration is effective for the CIST.
- In PVST mode, if you specify a VLAN list, the spanning tree configuration is effective for the specified VLANs. If you do not specify any VLAN, the spanning tree configuration is ineffective.

To set the spanning tree mode:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the spanning tree mode.	stp mode { stp rstp mstp pvst }	MSTP mode by default.

Configuring an MST region

Two or more spanning tree devices belong to the same MST region only if they are configured to have the same format selector (0 by default, not configurable), MST region name, MST region revision level, and the same VLAN-to-instance mapping entries in the MST region, and they are connected via a physical link.

Configuration restrictions and guidelines

- The configuration of MST region-related parameters, especially the VLAN-to-instance mapping table, will result in a new spanning tree calculation. To reduce the possibility of topology instability, the MST region configuration takes effect only after you activate it by using the **active region-configuration** command, or enable a spanning tree protocol by using the **stp enable** command in the case that the spanning tree protocol is disabled.
- The device in PVST mode automatically maps VLANs to MSTIs, and supports more MSTIs than in MSTP mode. When you change the spanning tree mode from PVST to MSTP, exceeding VLAN-to-instance mappings (arranged in ascending order of MSTI IDs) are silently deleted and cannot be recovered even if you change the spanning tree mode back. To prevent loss of mappings, do not manually configure VLAN-to-instance mappings in PVST mode.

Configuration procedure

To configure an MST region:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter MST region view.	stp region-configuration	N/A
3. Configure the MST region name.	region-name <i>name</i>	Optional. The MST region name is the MAC address by default.
4. Configure the VLAN-to-instance mapping table.	<ul style="list-style-type: none"> instance <i>instance-id</i> vlan <i>vlan-list</i> vlan-mapping modulo <i>modulo</i> 	Optional. Use either command. All VLANs in an MST region are mapped to the CIST (or MSTI 0) by default.
5. Configure the MSTP revision level of the MST region.	revision-level <i>level</i>	Optional. 0 by default.
6. Display the MST region configurations that are not activated yet.	check region-configuration	Optional.
7. Activate MST region configuration manually.	active region-configuration	N/A
8. Display the activated configuration information of the MST region.	display stp region-configuration [{ begin exclude include } <i>regular-expression</i>]	Optional. Available in any view

Configuring the root bridge or a secondary root bridge

You can have MSTP determine the root bridge of a spanning tree through MSTP calculation, or you can specify the current device as the root bridge or as a secondary root bridge using the commands that the system provides.

A device has independent roles in different spanning trees. It can act as the root bridge in one spanning tree and as a secondary root bridge in another. However, one device cannot be the root bridge and a secondary root bridge in the same spanning tree.

A spanning tree can have one root bridge only. If two or more devices are designated as the root bridge in a spanning tree at the same time, the device with the lowest MAC address wins.

When the root bridge of an instance fails or is shut down, the secondary root bridge (if you have specified one) can take over the role of the primary root bridge. However, if you specify a new primary root bridge for the instance then, the one you specify, not the secondary root bridge will become the root bridge. If you have specified multiple secondary root bridges for an instance, when the root bridge fails, the secondary root bridge with the lowest MAC address is selected as the new root bridge.

Configuration restrictions and guidelines

- You can specify one root bridge for each spanning tree, regardless of the device priority settings. Once you specify a device as the root bridge or a secondary root bridge, you cannot change its priority.

- You can configure the current device as the root bridge by setting the device priority to 0. For the device priority configuration, see "[Configuring the device priority](#)."

Configuring the current device as the root bridge of a specific spanning tree

To configure the current device as the root bridge of a specific spanning tree:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the current device as the root bridge.	<ul style="list-style-type: none"> In STP/RSTP mode: stp root primary In PVST mode: stp vlan <i>vlan-list</i> root primary In MSTP mode: stp [<i>instance instance-id</i>] root primary 	<p>Use one of the commands.</p> <p>By default, a device does not function as the root bridge.</p>

Configuring the current device as a secondary root bridge of a specific spanning tree

To configure the current device as a secondary root bridge of a specific spanning tree:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the current device as a secondary root bridge.	<ul style="list-style-type: none"> In STP/RSTP mode: stp root secondary In PVST mode: stp vlan <i>vlan-list</i> root secondary In MSTP mode: stp [<i>instance instance-id</i>] root secondary 	<p>Use one of the commands.</p> <p>By default, a device does not function as a secondary root bridge.</p>

Configuring the device priority

⚠ CAUTION:

- You cannot change the priority of a device after it is configured as the root bridge or as a secondary root bridge.
- During root bridge selection, if all devices in a spanning tree have the same priority, the one with the lowest MAC address will be selected as the root bridge of the spanning tree.

Device priority is a factor in spanning tree calculation. The priority of a device determines whether the device can be elected as the root bridge of a spanning tree. A lower numeric value indicates a higher priority. You can set the priority of a device to a low value to specify the device as the root bridge of the spanning tree. A spanning tree device can have different priorities in different MSTIs.

To configure the priority of a device in a specified MSTI:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the priority of the current device.	<ul style="list-style-type: none"> In STP/RSTP mode: stp priority <i>priority</i> In PVST mode: stp vlan <i>vlan-list</i> priority <i>priority</i> In MSTP mode: stp [instance <i>instance-id</i>] priority <i>priority</i> 	Use one of the commands. The default setting is 32768.

Configuring the maximum hops of an MST region

By setting the maximum hops of an MST region, you can restrict the region size. The maximum hops configured on the regional root bridge will be used as the maximum hops of the MST region.

Configuration BPDUs sent by the regional root bridge always have a hop count set to the maximum value. When a device receives this configuration BPDU, it decrements the hop count by 1, and uses the new hop count in the BPDUs that it propagates. When the hop count of a BPDU reaches 0, it is discarded by the device that received it. This prevents devices beyond the reach of the maximum hop from participate in spanning tree calculation, so the size of the MST region is limited.

Make this configuration on the root bridge only. All other devices in the MST region use the maximum hop value set for the root bridge.

To configure the maximum number of hops of an MST region:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the maximum hops of the MST region.	stp max-hops <i>hops</i>	20 by default.

Configuring the network diameter of a switched network

Any two terminal devices in a switched network are connected through a specific path composed of a series of devices. The network diameter is the number of devices on the path composed of the most devices. The network diameter is a parameter that indicates the network size. A bigger network diameter indicates a larger network size. Based on the network diameter you configured, the system automatically sets an optimal hello time, forward delay, and max age for the device.

To configure the network diameter of a switched network:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Configure the network diameter of the switched network.	<ul style="list-style-type: none"> In STP/RSTP/MSTP mode: stp bridge-diameter <i>diameter</i> In PVST mode: stp vlan <i>vlan-list</i> bridge-diameter <i>diameter</i> 	<p>Use one of the commands.</p> <p>The default setting is 7.</p>

NOTE:

- In STP/RSTP/MSTP mode, each MST region is considered as a device and the configured network diameter is effective only for the CIST (or the common root bridge), but not for MSTIs.
- In PVST mode, the network diameter configuration is effective on the root bridge only.

Configuring spanning tree timers

The following timers are used for spanning tree calculation:

- Forward delay

It is the delay time for port state transition. To prevent temporary loops on a network, the spanning tree sets an intermediate port state, the learning state, before it transitions from the discarding state to the forwarding state, and requires that the port transitions its state after a forward delay timer to make sure that the state transition of the local port keeps synchronized with the peer.
- Hello time

The device detects whether a link failure has occurred with the hello time interval. The spanning tree sends a configuration BPDU every hello time interval. If the device receives no configuration BPDUs within the hello time interval, it recalculates the spanning tree.
- Max age

In the CIST of an MSTP network or each VLAN of a PVST network, the device uses the max age parameter to determine whether a configuration BPDU received by a port has expired. If a port receives a configuration BPDU that has expired, that MSTI must be re-calculated. The max age timer is ineffective for MSTIs.

To avoid frequent network changes, be sure that the settings of the hello time, forward delay and max age timers meet the following formulas:

 - $2 \times (\text{forward delay} - 1 \text{ second}) \leq \text{max age}$
 - $\text{Max age} \leq 2 \times (\text{hello time} + 1 \text{ second})$

HP does not recommend you to manually set the spanning tree timers. Instead, you can specify the network diameter and let spanning tree protocols automatically calculate the timers based on the network diameter. If the network diameter uses the default value, the timers also use their default values.

Configure the timers on the root bridge only, and the timer settings on the root bridge apply to all devices on the entire switched network.

Configuration restrictions and guidelines

- The length of the forward delay timer is related to the network diameter of the switched network. The larger the network diameter is, the longer the forward delay time should be. If the forward delay timer is too short, temporary redundant paths might occur. If the forward delay timer is too long, network convergence might take a long time. HP recommends you to use the default setting.

- An appropriate hello time setting enables the device to quickly detect link failures on the network without using excessive network resources. If the hello time is too long, the device will mistake packet loss as a link failure and trigger a new spanning tree calculation process. If the hello time is too short, the device will frequently send the same configuration BPDUs, which adds the device burden and wastes network resources. HP recommends you to use the default setting.
- If the max age timer is too short, the device will frequently begin spanning tree calculation and might mistake network congestion as a link failure. If the max age timer is too long, the device might fail to quickly detect link failures and begin spanning tree calculations, reducing the auto-sensing capability of the network. HP recommends you to use the default setting.

Configuration procedure

To configure the spanning tree timers:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the forward delay timer.	<ul style="list-style-type: none"> • In STP/RSTP/MSTP mode: stp timer forward-delay time • In PVST mode: stp vlan vlan-list timer forward-delay time 	<p>Optional. Use one of the commands. The default setting is 15 seconds.</p>
3. Configure the hello timer.	<ul style="list-style-type: none"> • In STP/RSTP/MSTP mode: stp timer hello time • In PVST mode: stp vlan vlan-list timer hello time 	<p>Optional. Use one of the commands. The default setting is 2 seconds.</p>
4. Configure the max age timer.	<ul style="list-style-type: none"> • In STP/RSTP/MSTP mode: stp timer max-age time • In PVST mode: stp vlan vlan-list timer max-age time 	<p>Optional. Use one of the commands. The default setting is 20 seconds.</p>

Configuring the timeout factor

The timeout factor is a parameter used to decide the timeout time, in the following formula: Timeout time = timeout factor × 3 × hello time.

After the network topology is stabilized, each non-root-bridge device forwards configuration BPDUs to the downstream devices at the interval of hello time to determine whether any link is faulty. If a device does not receive a BPDU from the upstream device within nine times the hello time, it assumes that the upstream device has failed and starts a new spanning tree calculation process.

Sometimes a device might fail to receive a BPDU from the upstream device because the upstream device is busy. If a spanning tree calculation occurs, the calculation can fail and also waste network resources. In a stable network, you can prevent undesired spanning tree calculations by setting the timeout factor to 5, 6, or 7.

To configure the timeout factor:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Configure the timeout factor of the device.	stp timer-factor <i>factor</i>	3 by default.

Configuring the maximum port rate

The maximum rate of a port refers to the maximum number of BPDUs the port can send within each hello time. The maximum rate of a port is related to the physical status of the port and the network structure.

The higher the maximum port rate is, the more BPDUs will be sent within each hello time, and the more system resources will be used. By setting an appropriate maximum port rate, you can limit the rate at which the port sends BPDUs and prevent spanning tree protocols from using excessive network resources when the network becomes unstable. HP recommends you to use the default setting.

To configure the maximum rate of a port or a group of ports:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use one of the commands.
3. Configure the maximum rate of the ports.	stp transmit-limit <i>limit</i>	10 by default.

Configuring edge ports

If a port directly connects to a user terminal rather than another device or a shared LAN segment, this port is regarded as an edge port. When network topology change occurs, an edge port will not cause a temporary loop. Because a device does not determine whether a port is directly connected to a terminal, you must manually configure the port as an edge port. After that, the port can transition rapidly from the blocked state to the forwarding state.

Configuration restrictions and guidelines

- If BPDU guard is disabled, a port set as an edge port will become a non-edge port again if it receives a BPDU from another port. To restore the edge port, re-enable it.
- If a port directly connects to a user terminal, configure it as an edge port and enable BPDU guard for it. This enables the port to transition to the forwarding state quickly while ensuring network security.
- You cannot configure edge port settings and loop guard on a port at the same time.

Configuration procedure

To specify a port or a group of ports as edge port or ports:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use one of the commands.
3. Configure the current ports as edge ports.	stp edged-port enable	All ports are non-edge ports by default.

Configuring path costs of ports

Path cost is a parameter related to the rate of a port. On a spanning tree device, a port can have different path costs in different MSTIs. Setting appropriate path costs allows VLAN traffic flows to be forwarded along different physical links, achieving VLAN-based load balancing.

You can have the device automatically calculate the default path cost, or you can configure the path cost for ports.

Specifying a standard for the device to use when it calculates the default path cost

⚠ CAUTION:

If you change the standard that the device uses to calculate the default path costs, you restore the path costs to the default.

You can specify a standard for the device to use in automatic calculation for the default path cost. The device supports the following standards:

- **dot1d-1998**—The device calculates the default path cost for ports based on IEEE 802.1d-1998.
- **dot1t**—The device calculates the default path cost for ports based on IEEE 802.1t.
- **legacy**—The device calculates the default path cost for ports based on a private standard.

Table 14 shows the mappings between the link speed and the path cost.

Table 14 Mappings between the link speed and the path cost

Link speed	Port type	Path cost		
		IEEE 802.1d-1998	IEEE 802.1t	Private standard
0	N/A	65535	200,000,000	200,000
10 Mbps	Single port		2,000,000	2000
	Aggregate interface containing 2 Selected ports	100	1,000,000	1800
	Aggregate interface containing 3 Selected ports		666,666	1600

Link speed	Port type	Path cost		
		IEEE 802.1d-1998	IEEE 802.1t	Private standard
100 Mbps	Aggregate interface containing 4 Selected ports	19	500,000	1400
	Single port		200,000	200
	Aggregate interface containing 2 Selected ports		100,000	180
	Aggregate interface containing 3 Selected ports		66,666	160
	Aggregate interface containing 4 Selected ports		50,000	140
1000 Mbps	Single port	4	20,000	20
	Aggregate interface containing 2 Selected ports		10,000	18
	Aggregate interface containing 3 Selected ports		6666	16
	Aggregate interface containing 4 Selected ports		5000	14
10 Gbps	Single port	2	2000	2
	Aggregate interface containing 2 Selected ports		1000	1
	Aggregate interface containing 3 Selected ports		666	1
	Aggregate interface containing 4 Selected ports		500	1

Configuration restrictions and guidelines

- When it calculates path cost for an aggregate interface, IEEE 802.1t takes into account the number of Selected ports in its aggregation group, but IEEE 802.1d-1998 does not. The calculation formula of IEEE 802.1t is: Path cost = 200,000,000/link speed (in 100 kbps), where link speed is the sum of the link speed values of the Selected ports in the aggregation group.
- When multiple ports operate at a rate higher than 10 Gbps and the standard for default path cost calculation is **dot1d-1998** or **legacy**, the path cost of a single port or an aggregate interface takes the smallest value. As a result, the forwarding path selected might not be optimal. To solve this problem, use **dot1t** as the standard for default path cost calculation, or manually set the path cost for a port ([Configuring path costs of ports](#)).

Configuration procedure

To specify a standard for the device to use when it calculates the default path cost:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Specify a standard for the device to use when it calculates the default path costs of its ports.	stp pathcost-standard { dot1d-1998 dot1t legacy }	Optional. legacy by default.

Configuring path costs of ports

When the path cost of a port changes, the system re-calculates the role of the port and initiates a state transition.

To configure the path cost of ports:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use one of the commands.
3. Configure the path cost of the ports.	<ul style="list-style-type: none"> In STP/RSTP mode: stp cost <i>cost</i> In PVST mode: stp vlan <i>vlan-list</i> cost <i>cost</i> In MSTP mode: stp [instance <i>instance-id</i>] cost <i>cost</i> 	Use one of the commands. By default, the system automatically calculates the path cost of each port.

Configuration example

In MSTP mode, specify the device to calculate the default path costs of its ports by using IEEE 802.1d-1998, and set the path cost of GigabitEthernet 1/0/3 to 200 on MSTI 2.

```
<Sysname> system-view
[Sysname] stp pathcost-standard dot1d-1998
[Sysname] interface gigabitethernet 1/0/3
[Sysname-GigabitEthernet1/0/3] stp instance 2 cost 200
```

In PVST mode, specify the device to calculate the default path costs of its ports by using IEEE 802.1d-1998, and set the path cost of GigabitEthernet 1/0/3 to 2000 on VLANs 20 through 30.

```
<Sysname> system-view
[Sysname] stp mode pvst
[Sysname] stp pathcost-standard dot1d-1998
[Sysname] interface gigabitethernet 1/0/3
[Sysname-GigabitEthernet1/0/3] stp vlan 20 to 30 cost 2000
```

Configuring the port priority

When the priority of a port changes, MSTP re-calculates the role of the port and initiates a state transition. The priority of a port is an important factor in determining whether the port can be elected as the root port of a device. If all other conditions are the same, the port with the highest priority will be elected as the root port.

On a spanning tree device, a port can have different priorities and play different roles in different spanning trees, so that data of different VLANs can be propagated along different physical paths, implementing per-VLAN load balancing. You can set port priority values based on the actual networking requirements.

To configure the priority of a port or a group of ports:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none">Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i>Enter port group view: port-group manual <i>port-group-name</i>	Use one of the commands.
3. Configure the port priority.	<ul style="list-style-type: none">In STP/RSTP mode: stp port priority <i>priority</i>In PVST mode: stp vlan <i>vlan-list</i> port priority <i>priority</i>In MSTP mode: stp [instance <i>instance-id</i>] port priority <i>priority</i>	Use one of the commands. The default setting is 128.

Configuring the port link type

A point-to-point link directly connects two devices. If two root ports or designated ports are connected over a point-to-point link, they can rapidly transition to the forwarding state after a proposal-agreement handshake process.

Configuration restrictions and guidelines

- You can configure the link type as point-to-point for a Layer 2 aggregate interface or a port that operates in full duplex mode. HP recommends you to use the default setting and let the device to automatically detect the port link type.
- The **stp point-to-point force-false** or **stp point-to-point force-true** command configured on a port in MSTP or PVST mode is effective for all MSTIs or VLANs.
- If the physical link to which the port connects is not a point-to-point link but you set it to be one, the configuration might bring a temporary loop.

Configuration procedure

To configure the link type of a port or a group of ports:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use one of the commands.
3. Configure the port link type.	stp point-to-point { auto force-false force-true }	By default, the link type is auto where the port automatically detects the link type.

Configuring the mode a port uses to recognize/send MSTP packets

A port can receive/send MSTP packets in the following formats:

- **dot1s**—802.1s-compliant standard format
- **legacy**—Compatible format

By default, the packet format recognition mode of a port is **auto**. The port automatically distinguishes the two MSTP packet formats, and determines the format of packets that it will send based on the recognized format.

You can configure the MSTP packet format on a port. When operating in MSTP mode after the configuration, the port sends and receives only MSTP packets of the format that you have configured to communicate with devices that send packets of the same format.

MSTP provides MSTP packet format incompatibility guard. In MSTP mode, if a port is configured to recognize/send MSTP packets in a mode other than **auto**, and if it receives a packet in a format different from the specified type, the port becomes a designated port and remains in the discarding state to prevent the occurrence of a loop.

MSTP provides MSTP packet format frequent change guard. If a port receives MSTP packets of different formats frequently, the MSTP packet format configuration contains errors. If the port is operating in MSTP mode, it will be shut down for protection. Ports disabled in this way can be re-activated after a detection interval. For more information about the detection interval, see *Fundamentals Configuration Guide*.

To configure the MSTP packet format to be supported on a port or a group of ports:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use one of the commands.

Step	Command	Remarks
3. Configure the mode that the port uses to recognize/send MSTP packets.	stp compliance { auto dot1s legacy }	auto by default.

Enabling outputting port state transition information

In a large-scale spanning tree network, you can enable devices to output the port state transition information of all MSTIs or the specified MSTI in order to monitor the port states in real time.

To enable outputting port state transition information:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable outputting port state transition information.	<ul style="list-style-type: none"> In STP/RSTP mode: stp port-log instance 0 In PVST mode: stp port-log vlan <i>vlan-list</i> In MSTP mode: stp port-log instance { <i>instance-id</i> all } 	Use one of the commands. Enabled by default.

Enabling the spanning tree feature

You must enable the spanning tree feature for the device before any other spanning tree related configurations can take effect.

Configuration restrictions and guidelines

- To globally enable or disable the spanning tree feature (not for VLANs), use the **stp enable** command or **undo stp enable** command in system view. To enable or disable the spanning tree feature for specific VLANs, use the **stp vlan enable** command or **undo stp vlan enable** command.
- You can disable the spanning tree feature for certain ports with the **undo stp enable** command to exclude them from spanning tree calculation and save CPU resources of the device.
- In PVST mode, when you globally enable the spanning tree feature, the device automatically enables the spanning tree feature for the first n (which is the number of PVST instances that the switch supports and is 128 for the 5500 EI switch and 32 for the 5500 SI switch) of the existing VLANs by default. To enable the spanning tree feature for other VLANs, you must first disable the spanning tree feature for certain VLANs. This guideline does not apply if the number of existing VLANs on the switch does not exceed n .

Enabling the spanning tree feature (in STP/RSTP/MSTP mode)

In STP/RSTP/MSTP mode, make sure that the spanning tree feature is enabled globally and on the desired ports.

To enable the spanning tree feature in STP/RSTP/MSTP mode:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable the spanning tree feature globally.	stp enable	By default, the spanning tree feature is disabled globally.
3. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
4. Enable the spanning tree feature for the port or group of ports.	stp enable	Optional. By default, the spanning tree feature is enabled for all ports.

Enabling the spanning tree feature (in PVST mode)

In PVST mode, make sure that the spanning tree feature is enabled globally and on the desired VLANs and ports.

To enable the spanning tree feature in PVST mode:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Globally enable the spanning tree feature.	stp enable	By default, the spanning tree feature is disabled globally.
3. Enable the spanning tree feature on specific VLANs.	stp vlan <i>vlan-list</i> enable	By default, the spanning tree feature is enabled on VLANs.
4. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
5. Enable the spanning tree feature for the port or group of ports.	stp enable	Optional. By default, the spanning tree feature is enabled on all ports.

Performing mCheck

If a port on a device that is running MSTP, RSTP, or PVST connects to an STP device, this port automatically transitions to the STP-compatible mode. However, it cannot automatically transition back to the original mode under the following circumstances:

- The STP device is shut down or removed.
- The STP device transitions to the MSTP, RSTP, or PVST mode.

Suppose Device A running STP, Device B with no spanning tree feature enabled, and Device C running RSTP or MSTP are connected in order. Device B will transparently transmit the STP BPDUs, and the port on Device C and connecting to Device B will transition to the STP mode. After you enable the spanning tree feature on Device B, to run RSTP or MSTP between Device B and Device C, you must perform an mCheck operation on the ports interconnecting Device B and Device C, in addition to configuring the spanning tree to operate in RSTP or MSTP mode on Device B.

To forcibly transition the port to operate in the original mode, you can perform an mCheck operation. The following methods for performing mCheck produce the same result.

Performing mCheck globally

Step	Command
1. Enter system view.	system-view
2. Perform mCheck.	stp mcheck

Performing mCheck in interface view

Step	Command
1. Enter system view.	system-view
2. Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view.	interface <i>interface-type interface-number</i>
3. Perform mCheck.	stp mcheck

NOTE:

An mCheck operation takes effect on a device that operates in MSTP, RSTP, or PVST mode.

Configuring Digest Snooping

As defined in IEEE 802.1s, connected devices are in the same region only when their MST region-related configurations (region name, revision level, and VLAN-to-instance mappings) are identical. A spanning tree device identifies devices in the same MST region by determining the configuration ID in BPDU packets. The configuration ID includes the region name, revision level, and configuration digest, which is in 16-byte length and is the result calculated via the HMAC-MD5 algorithm based on VLAN-to-instance mappings.

Spanning tree implementations vary with vendors, and the configuration digests calculated using private keys is different, so devices of different vendors in the same MST region cannot communicate with each other.

To enable communication between an HP device and a third-party device, enable the Digest Snooping feature on the port that connects the HP device to the third-party device in the same MST region.

Configuration restrictions and guidelines

- Before you enable Digest Snooping, make sure that associated devices of different vendors are connected and run spanning tree protocols.

- With digest snooping enabled, in-the-same-region verification does not require comparison of configuration digest, so the VLAN-to-instance mappings must be the same on associated ports.
- With global Digest Snooping enabled, modification of VLAN-to-instance mappings and removal of the current region configuration via the **undo stp region-configuration** command are not allowed. You can modify only the region name and revision level.
- To make Digest Snooping take effect, you must enable it both globally and on associated ports. To make the configuration effective on all configured ports and while reducing impact on the network, enable Digest Snooping on all associated ports first and then globally.
- To prevent loops, do not enable Digest Snooping on MST region edge ports.
- HP recommends you to enable Digest Snooping first and then the spanning tree feature. To avoid causing traffic interruption, do not configure Digest Snooping when the network is already working well.

Configuration procedure

You can enable Digest Snooping only on the HP device that is connected to a third-party device that uses its private key to calculate the configuration digest.

To configure Digest Snooping:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> • Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> • Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Enable Digest Snooping on the interface or port group.	stp config-digest-snooping	Disabled by default.
4. Return to system view.	quit	N/A
5. Enable global Digest Snooping.	stp config-digest-snooping	Disabled by default.

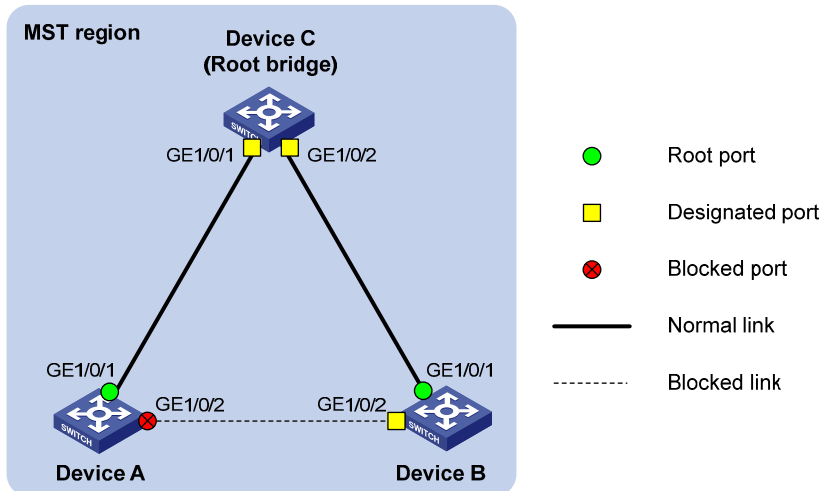
Digest Snooping configuration example

Network requirements

As shown in [Figure 23](#), Device A and Device B connect to Device C, which is a third-party device. All these devices are in the same region.

Enable Digest Snooping on the ports of Device A and Device B that connect to Device C, so that the three devices can communicate with one another.

Figure 23 Digest Snooping configuration



Configuration procedure

Enable Digest Snooping on GigabitEthernet 1/0/1 of Device A and enable global Digest Snooping on Device A.

```
<DeviceA> system-view
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] stp config-digest-snooping
[DeviceA-GigabitEthernet1/0/1] quit
[DeviceA] stp config-digest-snooping
```

Enable Digest Snooping on GigabitEthernet 1/0/1 of Device B and enable global Digest Snooping on Device B.

```
<DeviceB> system-view
[DeviceB] interface gigabitethernet 1/0/1
[DeviceB-GigabitEthernet1/0/1] stp config-digest-snooping
[DeviceB-GigabitEthernet1/0/1] quit
[DeviceB] stp config-digest-snooping
```

Configuring No Agreement Check

In RSTP and MSTP, the following types of messages are used for rapid state transition on designated ports:

- **Proposal**—Sent by designated ports to request rapid transition
- **Agreement**—Used to acknowledge rapid transition requests

Both RSTP and MSTP devices can perform rapid transition on a designated port only when the port receives an agreement packet from the downstream device. RSTP and MSTP devices have the following differences:

- For MSTP, the root port of the downstream device sends an agreement packet only after it receives an agreement packet from the upstream device.
- For RSTP, the downstream device sends an agreement packet regardless of whether an agreement packet from the upstream device is received.

Figure 24 Rapid state transition of an MSTP designated port

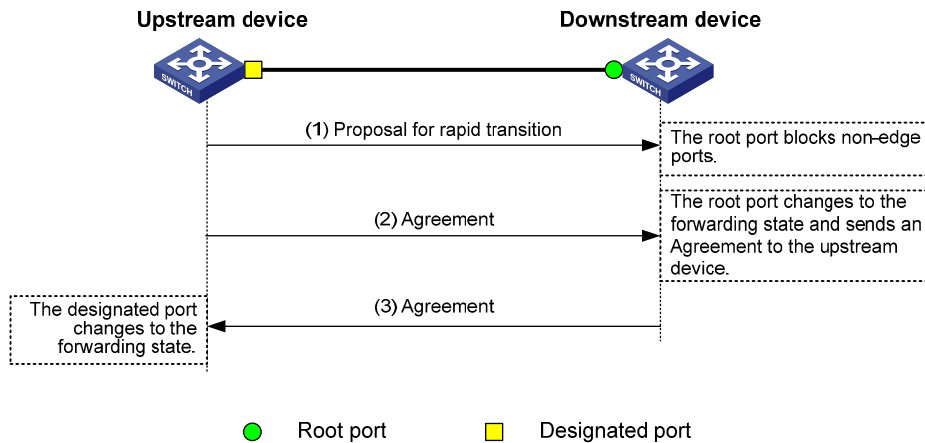
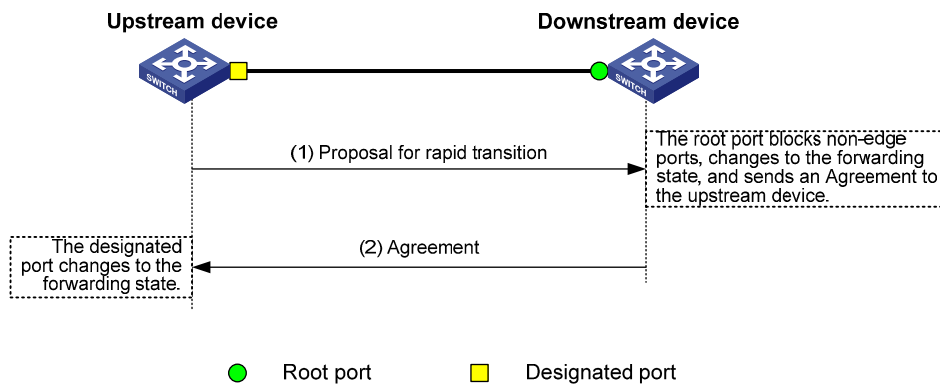


Figure 25 Rapid state transition of an RSTP designated port



If the upstream device is a third-party device, the rapid state transition implementation might be limited. For example, when the upstream device uses a rapid transition mechanism similar to that of RSTP, and the downstream device adopts MSTP and does not operate in RSTP mode, the root port on the downstream device receives no agreement packet from the upstream device and sends no agreement packets to the upstream device. As a result, the designated port of the upstream device fails to transit rapidly, and can only change to the forwarding state after a period twice the Forward Delay.

You can enable the No Agreement Check feature on the downstream device's port to enable the designated port of the upstream device to transit its state rapidly.

Configuration prerequisites

Before you configure the No Agreement Check function, complete the following tasks:

- Connect a device to a third-party upstream device that supports spanning tree protocols via a point-to-point link.
- Configure the same region name, revision level and VLAN-to-instance mappings on the two devices, assigning them to the same region.

Configuration procedure

To make the No Agreement Check feature take effect, enable it on the root port.

To configure No Agreement Check:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Enable No Agreement Check.	stp no-agreement-check	Disabled by default.

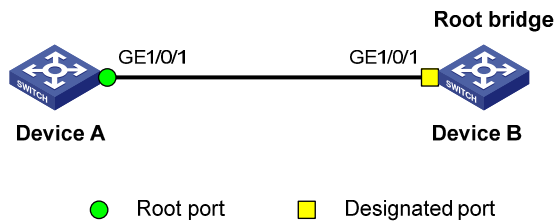
No Agreement Check configuration example

Network requirements

As shown in Figure 26:

- Device A connects to a third-party device that has a different spanning tree implementation. Both devices are in the same region.
- The third-party device (Device B) is the regional root bridge, and Device A is the downstream device.

Figure 26 Network diagram



Configuration procedure

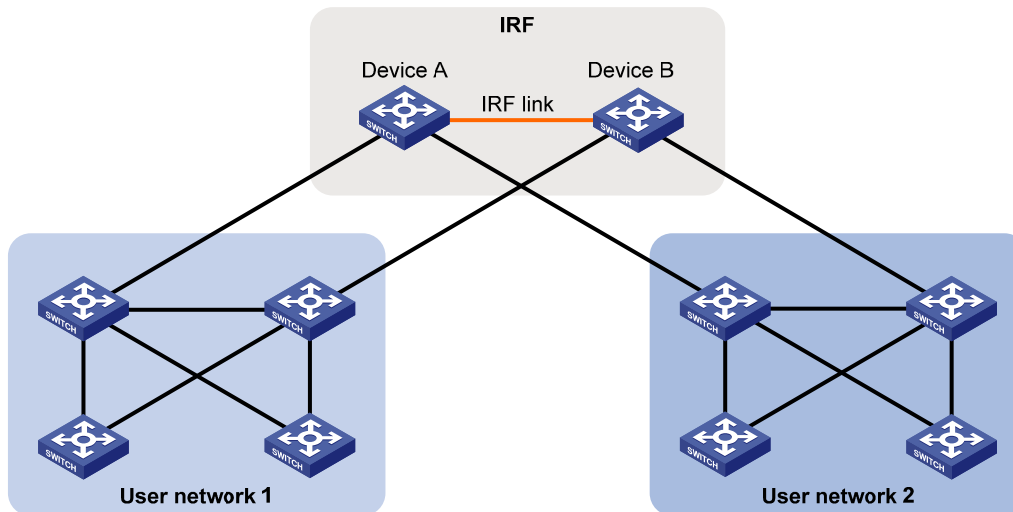
Enable No Agreement Check on GigabitEthernet 1/0/1 of Device A.

```
<DeviceA> system-view
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] stp no-agreement-check
```

Configuring TC snooping

Figure 27 shows a topology change (TC) snooping application scenario. Device A and Device B form an IRF fabric and do not have the spanning tree feature enabled. The IRF fabric connects to two user networks, in which all devices are enabled with the spanning tree feature. The user networks are dual-uplinked to the IRF fabric for high availability. The IRF fabric transparently transmits BPDUs in every user network.

Figure 27 TC snooping application scenario



In the network, the IRF fabric transparently transmits the received BPDUs and does not participate in spanning tree calculations. When a topology change occurs to the IRF fabric or user networks, the IRF fabric may need a long time to learn the correct MAC address table entries and ARP entries, resulting in long network disruption. To avoid the network disruption, you can enable TC snooping on the IRF fabric.

With TC snooping enabled, a device actively updates the MAC address table entries and ARP entries upon receiving TC-BPDUs, so that the device can normally forward the user traffic.

For more information about MAC address table entries, see "[Configuring the MAC address table.](#)"

For more information about ARP, see *Layer 3—IP Services Configuration Guide*.

Configuration restrictions and guidelines

- TC snooping and STP are mutually exclusive. You must globally disable the spanning tree feature before enable TC snooping.
- TC snooping does not take effect on the ports on which BPDU tunneling is enabled for spanning tree protocols. For more information about BPDU tunneling, see "[Configuring BPDU tunneling.](#)"
- TC snooping does not support PVST TC-BPDUs. As a result, TC snooping does not take effect in a PVST network.

Configuration procedure

To configure TC snooping:

Step	Command	Description
1. Enter system view.	system-view	N/A
2. Globally disable the spanning tree feature.	undo stp enable	By default, the spanning tree feature is disabled globally.
3. Enable TC snooping.	stp tc-snooping	Disabled by default.

Configuring protection functions

A spanning tree device supports the following protection functions:

- BPDU guard
- Root guard
- Loop guard
- TC-BPDU guard
- BPDU drop

Configuration prerequisites

The spanning tree feature has been correctly configured on the device.

Enabling BPDU guard

For access layer devices, the access ports can directly connect to the user terminals (such as PCs) or file servers. The access ports are configured as edge ports to allow rapid transition. When these ports receive configuration BPDUs, the system automatically sets the ports as non-edge ports and starts a new spanning tree calculation process. This causes a change of network topology. Under normal conditions, these ports should not receive configuration BPDUs. However, if someone forges configuration BPDUs maliciously to attack the devices, the network will become unstable.

The spanning tree protocol provides the BPDU guard function to protect the system against such attacks. With the BPDU guard function enabled on the devices, when edge ports receive configuration BPDUs, the system closes these ports and notifies the NMS that these ports have been closed by the spanning tree protocol. The device will reactivate the closed ports after a detection interval. For more information about this detection interval, see *Fundamentals Configuration Guide*.

Configure BPDU guard on a device with edge ports configured.

To enable BPDU guard:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable the BPDU guard function for the device.	stp bpdu-protection	Disabled by default.

NOTE:

BPDU guard does not take effect on loopback-testing-enabled ports. For more information about loopback testing, see "[Configuring Ethernet interfaces](#)."

Enabling root guard

The root bridge and secondary root bridge of a spanning tree should be located in the same MST region. Especially for the CIST, the root bridge and secondary root bridge are put in a high-bandwidth core region during network design. However, due to possible configuration errors or malicious attacks in the network, the legal root bridge might receive a configuration BPDU with a higher priority. Another device will supersede the current legal root bridge, causing an undesired change of the network topology. The

traffic that should go over high-speed links is switched to low-speed links, resulting in network congestion.

To prevent this situation, MSTP provides the root guard function. If the root guard function is enabled on a port of a root bridge, this port plays the role of designated port on all MSTIs. After this port receives a configuration BPDU with a higher priority from an MSTI, it immediately sets that port to the listening state in the MSTI, without forwarding the packet. This is equivalent to disconnecting the link connected with this port in the MSTI. If the port receives no BPDUs with a higher priority within twice the forwarding delay, it reverts to its original state.

Configure root guard on a designated port.

To enable root guard:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Enable the root guard function for the port(s).	stp root-protection	Disabled by default.

NOTE:

You cannot configure root guard and loop guard on a port at the same time.

Enabling loop guard

A device that keeps receiving BPDUs from the upstream device can maintain the state of the root port and blocked ports. However, link congestion or unidirectional link failures might cause these ports to fail to receive BPDUs from the upstream devices. The device will reselect the port roles: Those ports in forwarding state that failed to receive upstream BPDUs will become designated ports, and the blocked ports will transition to the forwarding state, resulting in loops in the switched network. The loop guard function can suppress the occurrence of such loops.

The initial state of a loop guard-enabled port is discarding in every MSTI. When the port receives BPDUs, its state transitions normally. Otherwise, it stays in the discarding state to prevent temporary loops.

Configure loop guard on the root port and alternate ports of a device.

To enable loop guard:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Enable the loop guard function for the ports.	stp loop-protection	Disabled by default.

NOTE:

- Do not enable loop guard on a port that connects user terminals. Otherwise, the port will stay in the discarding state in all MSTIs because it cannot receive BPDUs.
- You cannot configure edge port settings and loop guard, or configure root guard and loop guard on a port at the same time.

Enabling TC-BPDU guard

When a switch receives topology change (TC) BPDUs (the BPDUs that notify devices of topology changes), the switch flushes its forwarding address entries. If someone forges TC-BPDUs to attack the switch, the switch will receive a large number of TC-BPDUs within a short time and be busy with forwarding address entry flushing. This affects network stability.

With the TC-BPDU guard function, you can set the maximum number of immediate forwarding address entry flushes that the device can perform every a specified period of time (10 seconds). For TC-BPDUs received in excess of the limit, the device performs a forwarding address entry flush when the time period expires. This prevents frequent flushing of forwarding address entries.

To enable TC-BPDU guard:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable the TC-BPDU guard function.	stp tc-protection enable	Optional. Enabled by default.
3. Configure the maximum number of forwarding address entry flushes that the device can perform every 10 seconds.	stp tc-protection threshold <i>number</i>	Optional. 6 by default.

NOTE:

HP does not recommend you disable this feature.

Enabling BPDU drop

In a spanning tree network, after receiving BPDUs, the device performs STP calculation according to the received BPDUs and forwards received BPDUs to other devices in the network. This allows malicious attackers to attack the network by forging BPDUs. By continuously sending forged BPDUs, they can make

all the devices in the network perform STP calculations all the time. As a result, problems such as CPU overload and BPDU protocol status errors occur.

To avoid this problem, you can enable BPDU drop on ports. A BPDU drop-enabled port does not receive any BPDUs and is invulnerable to forged BPDU attacks.

To enable BPDU drop on an Ethernet interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable BPDU drop on the current interface.	bpdu-drop any	Disabled by default.

NOTE:

Because a port with BPDU drop enabled also drops the received 802.1X packets, do not enable BPDU drop and 802.1X on a port at the same time. For more information about 802.1X, see *Security Configuration Guide*.

Displaying and maintaining the spanning tree

Task	Command	Remarks
Display information about ports blocked by spanning tree protection functions.	display stp abnormal-port [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display BPDU statistics on ports.	display stp bpdu-statistics [interface <i>interface-type interface-number</i> [instance <i>instance-id</i>]] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about ports shut down by spanning tree protection functions.	display stp down-port [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the historical information of port role calculation for the specified MSTI or all MSTIs.	display stp [instance <i>instance-id</i> vlan <i>vlan-id</i>] history [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the statistics of TC/TCN BPDUs sent and received by all ports in the specified MSTI or all MSTIs.	display stp [instance <i>instance-id</i> vlan <i>vlan-id</i>] tc [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the spanning tree status and statistics.	display stp [instance <i>instance-id</i> vlan <i>vlan-id</i>] [interface <i>interface-list</i> slot <i>slot-number</i>] [brief] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the MST region configuration information that has taken effect.	display stp region-configuration [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the root bridge information of all MSTIs.	display stp root [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Clear the spanning tree statistics.	<code>reset stp [interface interface-list]</code>	Available in user view

Spanning tree configuration examples

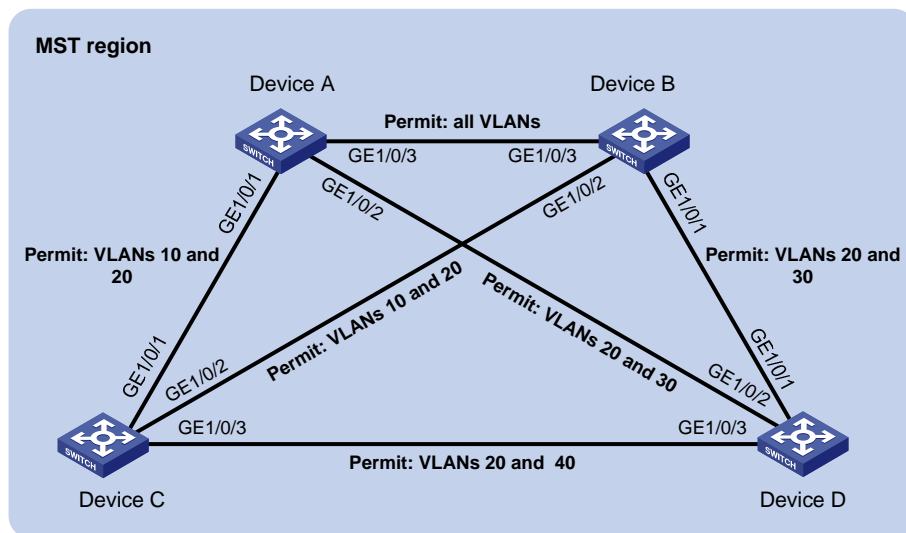
MSTP configuration example

Network requirements

As shown in [Figure 28](#):

- All devices on the network are in the same MST region. Device A and Device B work at the distribution layer. Device C and Device D work at the access layer.
- Configure MSTP so that packets of different VLANs are forwarded along different spanning trees: Packets of VLAN 10 are forwarded along MSTI 1, those of VLAN 30 are forwarded along MSTI 3, those of VLAN 40 are forwarded along MSTI 4, and those of VLAN 20 are forwarded along MSTI 0.
- VLAN 10 and VLAN 30 are terminated on the distribution layer devices, and VLAN 40 is terminated on the access layer devices. The root bridges of MSTI 1 and MSTI 3 are Device A and Device B, respectively, and the root bridge of MSTI 4 is Device C.

Figure 28 Network diagram



Configuration procedure

1. Configure VLANs and VLAN member ports (Details not shown.).
Create VLAN 10, VLAN 20, and VLAN 30 on Device A and Device B, respectively, VLAN 10, VLAN 20, and VLAN 40 on Device C, and VLAN 20, VLAN 30, and VLAN 40 on Device D. Configure the ports on these devices as trunk ports and assign them to related VLANs.
2. Configure Device A:
Enter MST region view; configure the MST region name as **example**; map VLAN 10, VLAN 30, and VLAN 40 to MSTI 1, MSTI 3, and MSTI 4, respectively; configure the revision level of the MST region as 0.

```

<DeviceA> system-view
[DeviceA] stp region-configuration
[DeviceA-mst-region] region-name example
[DeviceA-mst-region] instance 1 vlan 10
[DeviceA-mst-region] instance 3 vlan 30
[DeviceA-mst-region] instance 4 vlan 40
[DeviceA-mst-region] revision-level 0
# Activate MST region configuration.
[DeviceA-mst-region] active region-configuration
[DeviceA-mst-region] quit
# Specify the current device as the root bridge of MSTI 1.
[DeviceA] stp instance 1 root primary
# Enable the spanning tree feature globally.
[DeviceA] stp enable

```

3. Configure Device B:

Enter MST region view, configure the MST region name as **example**, map VLAN 10, VLAN 30, and VLAN 40 to MSTI 1, MSTI 3, and MSTI 4, respectively, and configure the revision level of the MST region as 0.

```

<DeviceB> system-view
[DeviceB] stp region-configuration
[DeviceB-mst-region] region-name example
[DeviceB-mst-region] instance 1 vlan 10
[DeviceB-mst-region] instance 3 vlan 30
[DeviceB-mst-region] instance 4 vlan 40
[DeviceB-mst-region] revision-level 0
# Activate MST region configuration.
[DeviceB-mst-region] active region-configuration
[DeviceB-mst-region] quit
# Specify the current device as the root bridge of MSTI 3.
[DeviceB] stp instance 3 root primary
# Enable the spanning tree feature globally.
[DeviceB] stp enable

```

4. Configure Device C:

Enter MST region view, configure the MST region name as **example**, map VLAN 10, VLAN 30, and VLAN 40 to MSTI 1, MSTI 3, and MSTI 4, respectively, and configure the revision level of the MST region as 0.

```

<DeviceC> system-view
[DeviceC] stp region-configuration
[DeviceC-mst-region] region-name example
[DeviceC-mst-region] instance 1 vlan 10
[DeviceC-mst-region] instance 3 vlan 30
[DeviceC-mst-region] instance 4 vlan 40
[DeviceC-mst-region] revision-level 0
# Activate MST region configuration.
[DeviceC-mst-region] active region-configuration
[DeviceC-mst-region] quit

```

Specify the current device as the root bridge of MSTI 4.

```
[DeviceC] stp instance 4 root primary
```

Enable the spanning tree feature globally.

```
[DeviceC] stp enable
```

5. Configure Device D:

Enter MST region view, configure the MST region name as **example**, map VLAN 10, VLAN 30, and VLAN 40 to MSTI 1, MSTI 3, and MSTI 4, respectively, and configure the revision level of the MST region as 0.

```
<DeviceD> system-view
```

```
[DeviceD] stp region-configuration
```

```
[DeviceD-mst-region] region-name example
```

```
[DeviceD-mst-region] instance 1 vlan 10
```

```
[DeviceD-mst-region] instance 3 vlan 30
```

```
[DeviceD-mst-region] instance 4 vlan 40
```

```
[DeviceD-mst-region] revision-level 0
```

Activate MST region configuration.

```
[DeviceD-mst-region] active region-configuration
```

```
[DeviceD-mst-region] quit
```

Enable the spanning tree feature globally.

```
[DeviceD] stp enable
```

6. Verify the configurations:

You can use the **display stp brief** command to display brief spanning tree information on each device after the network is stable.

Display brief spanning tree information on Device A.

```
[DeviceA] display stp brief
```

MSTID	Port	Role	STP State	Protection
0	GigabitEthernet1/0/1	ALTE	DISCARDING	NONE
0	GigabitEthernet1/0/2	DESI	FORWARDING	NONE
0	GigabitEthernet1/0/3	ROOT	FORWARDING	NONE
1	GigabitEthernet1/0/1	DESI	FORWARDING	NONE
1	GigabitEthernet1/0/3	DESI	FORWARDING	NONE
3	GigabitEthernet1/0/2	DESI	FORWARDING	NONE
3	GigabitEthernet1/0/3	ROOT	FORWARDING	NONE

Display brief spanning tree information on Device B.

```
[DeviceB] display stp brief
```

MSTID	Port	Role	STP State	Protection
0	GigabitEthernet1/0/1	DESI	FORWARDING	NONE
0	GigabitEthernet1/0/2	DESI	FORWARDING	NONE
0	GigabitEthernet1/0/3	DESI	FORWARDING	NONE
1	GigabitEthernet1/0/2	DESI	FORWARDING	NONE
1	GigabitEthernet1/0/3	ROOT	FORWARDING	NONE
3	GigabitEthernet1/0/1	DESI	FORWARDING	NONE
3	GigabitEthernet1/0/3	DESI	FORWARDING	NONE

Display brief spanning tree information on Device C.

```
[DeviceC] display stp brief
```

MSTID	Port	Role	STP State	Protection
-------	------	------	-----------	------------

```

0      GigabitEthernet1/0/1      DESI FORWARDING NONE
0      GigabitEthernet1/0/2      ROOT FORWARDING NONE
0      GigabitEthernet1/0/3      DESI FORWARDING NONE
1      GigabitEthernet1/0/1      ROOT FORWARDING NONE
1      GigabitEthernet1/0/2      ALTE DISCARDING NONE
4      GigabitEthernet1/0/3      DESI FORWARDING NONE

```

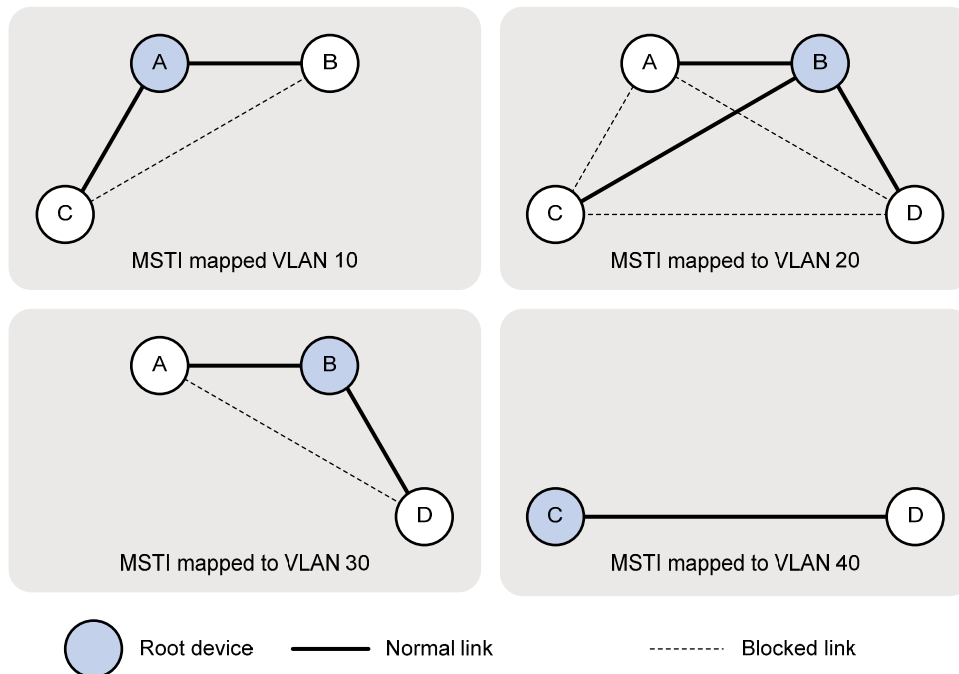
Display brief spanning tree information on Device D.

```
[DeviceD] display stp brief
```

MSTID	Port	Role	STP State	Protection
0	GigabitEthernet1/0/1	ROOT	FORWARDING	NONE
0	GigabitEthernet1/0/2	ALTE	DISCARDING	NONE
0	GigabitEthernet1/0/3	ALTE	DISCARDING	NONE
3	GigabitEthernet1/0/1	ROOT	FORWARDING	NONE
3	GigabitEthernet1/0/2	ALTE	DISCARDING	NONE
4	GigabitEthernet1/0/3	ROOT	FORWARDING	NONE

Based on the output, you can draw the MSTI mapped to each VLAN, as shown in [Figure 29](#).

Figure 29 MSTIs mapped to different VLANs



PVST configuration example

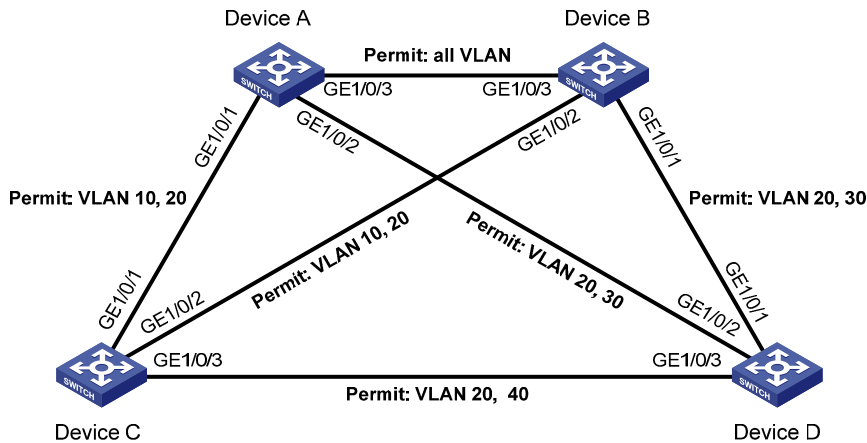
Network requirements

As shown in [Figure 30](#):

- Device A and Device B work at the distribution layer. Device C and Device D work at the access layer.
- Configure PVST so that packets of different VLANs are forwarded along different spanning trees.

- VLAN 10, VLAN 20, and VLAN 30 are terminated on the distribution layer devices, and VLAN 40 is terminated on the access layer devices. The root bridge of VLAN 10 and VLAN 20 is Device A, that of VLAN 30 is Device B, and that of VLAN 40 is Device C.

Figure 30 Network diagram



Configuration procedure

1. Configure VLANs and VLAN member ports. (Details not shown.)
Create VLAN 10, VLAN 20, and VLAN 30 on Device A and Device B, respectively, VLAN 10, VLAN 20, and VLAN 40 on Device C, and VLAN 20, VLAN 30, and VLAN 40 on Device D. Configure the ports on these devices as trunk ports and assign them to related VLANs.
2. Configure Device A:
 - # Set the spanning tree mode to PVST.

```
<DeviceA> system-view
[DeviceA] stp mode pvst
```
- # Specify the device as the root bridge of VLAN 10 and VLAN 20.


```
[DeviceA] stp vlan 10 20 root primary
```
- # Enable the spanning tree feature globally and for VLANs 10, 20, and 30.


```
[DeviceA] stp enable
[DeviceA] stp vlan 10 20 30 enable
```
3. Configure Device B:
 - # Set the spanning tree mode to PVST.

```
<DeviceB> system-view
[DeviceB] stp mode pvst
```
- # Specify the device as the root bridge of VLAN 30.


```
[DeviceB] stp vlan 30 root primary
```
- # Enable the spanning tree feature globally and for VLANs 10, 20, and 30.


```
[DeviceB] stp enable
[DeviceB] stp vlan 10 20 30 enable
```
4. Configure Device C:
 - # Set the spanning tree mode to PVST.

```
<DeviceC> system-view
[DeviceC] stp mode pvst
```

Specify the current device as the root bridge of VLAN 40.

```
[DeviceC] stp vlan 40 root primary
```

Enable the spanning tree feature globally and for VLANs 10, 20, and 40.

```
[DeviceC] stp enable
```

```
[DeviceC] stp vlan 10 20 40 enable
```

5. Configure Device D:

Set the spanning tree mode to PVST.

```
<DeviceD> system-view
```

```
[DeviceD] stp mode pvst
```

Enable the spanning tree feature globally and for VLANs 20, 30, and 40.

```
[DeviceD] stp enable
```

```
[DeviceD] stp vlan 20 30 40 enable
```

6. Verify the configurations:

You can use the **display stp brief** command to display brief spanning tree information on each device after the network is stable.

Display brief spanning tree information on Device A.

```
[DeviceA] display stp brief
```

VLAN	Port	Role	STP State	Protection
10	GigabitEthernet1/0/1	DESI	DISCARDING	NONE
10	GigabitEthernet1/0/3	DESI	FORWARDING	NONE
20	GigabitEthernet1/0/1	DESI	FORWARDING	NONE
20	GigabitEthernet1/0/2	DESI	FORWARDING	NONE
20	GigabitEthernet1/0/3	DESI	FORWARDING	NONE
30	GigabitEthernet1/0/2	DESI	FORWARDING	NONE
30	GigabitEthernet1/0/3	ROOT	FORWARDING	NONE

Display brief spanning tree information on Device B.

```
[DeviceB] display stp brief
```

VLAN	Port	Role	STP State	Protection
10	GigabitEthernet1/0/2	DESI	FORWARDING	NONE
10	GigabitEthernet1/0/3	ROOT	FORWARDING	NONE
20	GigabitEthernet1/0/1	DESI	FORWARDING	NONE
20	GigabitEthernet1/0/2	DESI	FORWARDING	NONE
20	GigabitEthernet1/0/3	ROOT	FORWARDING	NONE
30	GigabitEthernet1/0/1	DESI	FORWARDING	NONE
30	GigabitEthernet1/0/3	DESI	FORWARDING	NONE

Display brief spanning tree information on Device C.

```
[DeviceC] display stp brief
```

VLAN	Port	Role	STP State	Protection
10	GigabitEthernet1/0/1	ROOT	FORWARDING	NONE
10	GigabitEthernet1/0/2	ALTE	FORWARDING	NONE
20	GigabitEthernet1/0/1	ROOT	FORWARDING	NONE
20	GigabitEthernet1/0/2	ALTE	FORWARDING	NONE
20	GigabitEthernet1/0/3	DESI	DISCARDING	NONE
40	GigabitEthernet1/0/3	DESI	FORWARDING	NONE

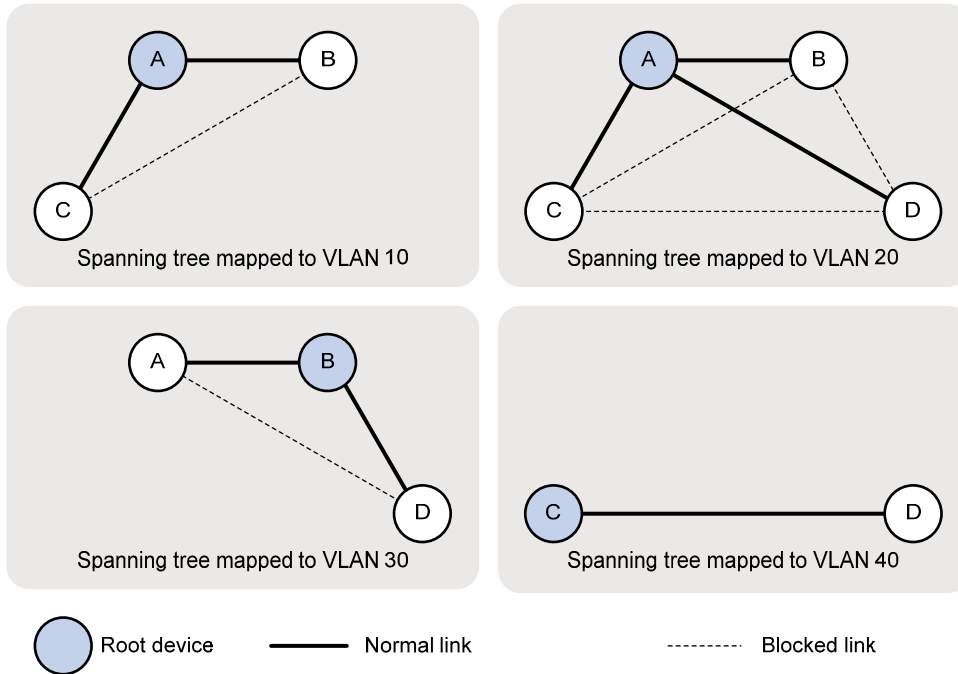
Display brief spanning tree information on Device D.

```
[DeviceD] display stp brief
```

VLAN	Port	Role	STP State	Protection
20	GigabitEthernet1/0/1	ALTE	FORWARDING	NONE
20	GigabitEthernet1/0/2	ROOT	DISCARDING	NONE
20	GigabitEthernet1/0/3	ALTE	DISCARDING	NONE
30	GigabitEthernet1/0/1	ROOT	FORWARDING	NONE
30	GigabitEthernet1/0/2	ALTE	DISCARDING	NONE
40	GigabitEthernet1/0/3	ROOT	FORWARDING	NONE

Based on the output, you can draw the spanning tree mapped to each VLAN, as shown in [Figure 31](#).

Figure 31 Spanning trees mapped to different VLANs



Configuring BPDU tunneling

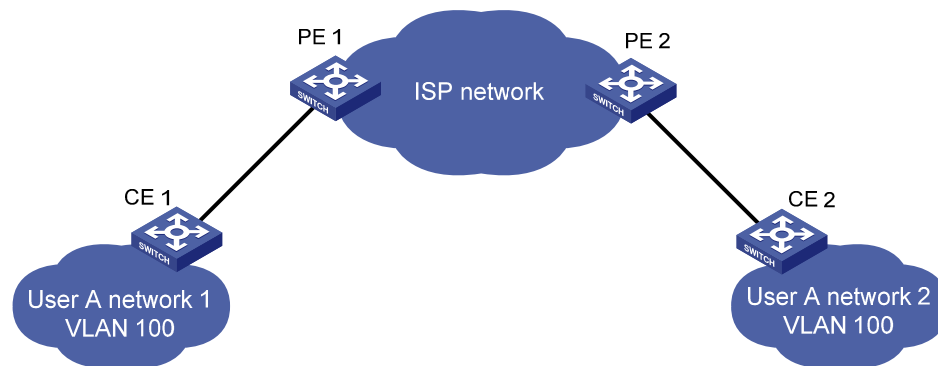
Overview

As a Layer 2 tunneling technology, BPDU tunneling enables Layer 2 protocol packets from geographically dispersed customer networks to be transparently transmitted over specific tunnels across a service provider network.

Background

Dedicated lines are used in a service provider network to build user-specific Layer 2 networks. As a result, a user network consists of parts located at different sides of the service provider network. As shown in Figure 32, the devices for User A are CE 1 and CE 2, both of which belong to VLAN 100. User A's network is divided into network 1 and network 2, which are connected by the service provider network. When a Layer 2 protocol (for example, STP) runs on both network 1 and network 2, the Layer 2 protocol packets must be transmitted over the service provider network to implement Layer 2 protocol calculation (for example, spanning tree calculation). When receiving a Layer 2 protocol packet, the PEs cannot determine whether the packet is from the user network or the service provider network, and must deliver the packet to the CPU for processing. In this case, the Layer 2 protocol calculation in User A's network is mixed with that in the service provider network, and the user network cannot implement independent Layer 2 protocol calculation.

Figure 32 BPDU tunneling application scenario



BPDU tunneling addresses this problem. With BPDU tunneling, Layer 2 protocol packets from customer networks can be transparently transmitted over the service provider network in the following workflow:

1. After receiving a Layer 2 protocol packet from CE 1, PE 1 encapsulates the packet, replaces its destination MAC address with a specific multicast MAC address, and forwards the packet to the service provider network.
2. The encapsulated Layer 2 protocol packet (called bridge protocol data unit, BPDU) is forwarded to PE 2 at the other end of the service provider network, which de-encapsulates the packet, restores the original destination MAC address of the packet, and then sends the packet to CE 2.

HP devices support BPDU tunneling for the following protocols:

- Cisco Discovery Protocol (CDP)
- Device Link Detection Protocol (DLDP)

- Ethernet Operation, Administration and Maintenance (EOAM)
- GARP VLAN Registration Protocol (GVRP)
- HW Group Management Protocol (HGMP)
- Link Aggregation Control Protocol (LACP)
- Link Layer Discovery Protocol (LLDP)
- Port Aggregation Protocol (PAGP)
- Per VLAN Spanning Tree (PVST)
- Spanning Tree Protocol (STP)
- Unidirectional Link Direction (UDLD)
- VLAN Trunking Protocol (VTP)

BPDU tunneling implementation

The BPDU tunneling implementations for different protocols are all similar. This section uses the Spanning Tree Protocol (STP) to describe how to implement BPDU tunneling.

This document uses the term *STP* in a broad sense. It includes STP, RSTP, and MSTP.

STP calculates the topology of a network by transmitting BPDUs among devices in the network. For more information, see "[Configuring spanning tree protocols.](#)"

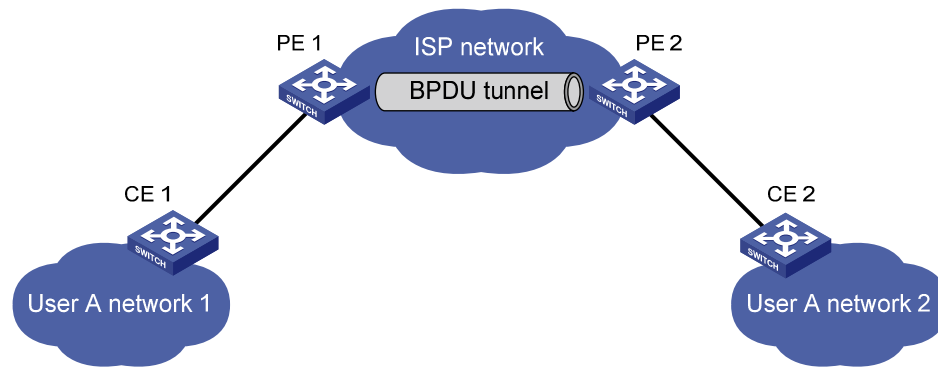
To avoid loops in your network, you can enable STP on your devices. When the topology changes at one side of the customer network, devices at that side of the customer network send BPDUs to devices on the other side of the customer network to ensure consistent spanning tree calculation in the entire customer network. However, because BPDUs are Layer 2 multicast frames, all STP-enabled devices, both in the customer network and in the service provider network, can receive and process these BPDUs. In this case, neither the service provider network nor the customer network can correctly calculate its independent spanning tree.

BPDU tunneling allows each network to calculate an independent spanning tree with STP.

BPDU tunneling delivers the following benefits:

- BPDUs can be transparently transmitted. BPDUs of one customer network can be broadcast in a specific VLAN across the service provider network, allowing that customer's geographically dispersed networks to implement consistent spanning tree calculation across the service provider network.
- BPDUs of different customer networks can be confined within different VLANs for transmission on the service provider network. This enables each customer network to perform independent spanning tree calculation.

Figure 33 BPDU tunneling implementation



The upper section of [Figure 33](#) represents the service provider network (ISP network). The lower section, including User A network 1 and User A network 2, represents the customer networks. Enabling BPDU tunneling on edge devices (PE 1 and PE 2) in the service provider network allows BPDUs of User A network 1 and User A network 2 to be transparently transmitted through the service provider network. This ensures consistent spanning tree calculation throughout User A network, without affecting the spanning tree calculation of the service provider network.

Assume that a BPDU is sent from User A network 1 to User A network 2. The BPDU is sent by using the following workflow.

1. At the ingress of the service provider network, PE 1 changes the destination MAC address of the BPDU from 0x0180-C200-0000 to a special multicast MAC address, 0x010F-E200-0003 (the default multicast MAC address), for example. In the service provider network, the modified BPDU is forwarded as a data packet in the VLAN assigned to User A.
2. At the egress of the service provider network, PE 2 recognizes the BPDU with the destination MAC address 0x010F-E200-0003, restores its original destination MAC address 0x0180-C200-0000, and then sends the BPDU to CE 2.

NOTE:

Through configuration, make sure that the VLAN tags carried in BPDUs are neither changed nor removed during the transparent transmission in the service provider network. Otherwise, the devices in the service provider network will fail to transparently transmit the customer network BPDUs correctly.

Enabling BPDU tunneling

Configuration prerequisites

Before configuring BPDU tunneling for a protocol, perform the following tasks:

- Enable the protocol in the customer network.
- Assign the port on which you want to enable BPDU tunneling on the PE device and the connected port on the CE device to the same VLAN.
- Configure ports that connect network devices in the service provider network as trunk ports that allow packets of any VLAN to pass through.

Configuration restrictions and guidelines

- Settings made in Layer 2 Ethernet interface view or Layer 2 aggregate interface view take effect only on the current port. Settings made in port group view take effect on all ports in the port group.
- Before you enable BPDU tunneling for DLDAP, EOAM, GVRP, HGMP, LLDP, or STP on a port, disable the protocol on the port first.
- Because PVST is a special STP protocol, you must do two things before you enable BPDU tunneling for PVST on a port: first, disable STP; second, enable BPDU tunneling for STP on the port.
- Do not enable BPDU tunneling for DLDAP, EOAM, LACP, LLDP, PAGP, or UDLD on the member port of a Layer 2 aggregation group.

Enabling BPDU tunneling

You can enable BPDU tunneling for different protocols in different views.

Enabling BPDU tunneling for a protocol in Layer 2 Ethernet interface view or port group view

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view or port group view.	<ul style="list-style-type: none">• Enter Layer 2 Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i>• Enter port group view: port-group manual <i>port-group-name</i>	Use either command.
3. Enable BPDU tunneling for a protocol.	bpdu-tunnel dot1q { cdp dldp eoam gvrp hgmp lacp lldp pagp pvst stp udld vtp }	Disabled by default.

Enabling BPDU tunneling for a protocol in Layer 2 aggregate interface view

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 aggregate interface view.	interface bridge-aggregation <i>interface-number</i>	N/A
3. Enable BPDU tunneling for a protocol on the Layer 2 aggregate interface.	bpdu-tunnel dot1q { cdp gvrp hgmp pvst stp vtp }	Disabled by default.

Configuring destination multicast MAC address for BPDUs

By default, the destination multicast MAC address for BPDUs is 0x010F-E200-0003. You can change it to 0x0100-0CCD-CDD0, 0x0100-0CCD-CDD1, or 0x0100-0CCD-CDD2.

To configure destination multicast MAC address for BPDUs:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the destination multicast MAC address for BPDUs.	bpdu-tunnel tunnel-dmac <i>mac-address</i>	Optional. 0x010F-E200-0003 by default.

NOTE:

For BPDUs to be recognized, the destination multicast MAC addresses configured for BPDU tunneling must be the same on the edge devices on the service provider network.

BPDU tunneling configuration examples

BPDU tunneling for STP configuration example

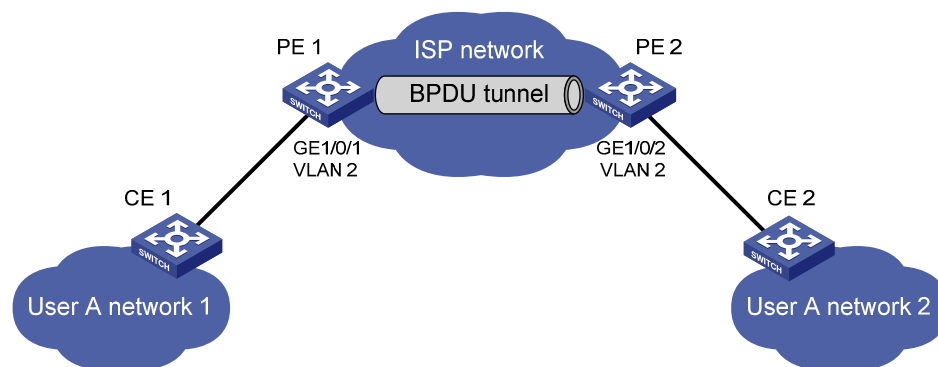
Network requirements

As shown in Figure 34:

- CE 1 and CE 2 are edge devices on the geographically dispersed network of User A; PE 1 and PE 2 are edge devices on the service provider network.
- All ports that connect service provider devices and customer devices are access ports and belong to VLAN 2. All ports that interconnect service provider devices are trunk ports and allow packets of any VLAN to pass through.
- MSTP is enabled on User A's network.

After the configuration, CE 1 and CE 2 must implement consistent spanning tree calculation across the service provider network, and the destination multicast MAC address carried in BPDUs must be 0x0100-0CCD-CDD0.

Figure 34 Network diagram



Configuration procedure

1. Configure PE 1:

Configure the destination multicast MAC address for BPDUs as 0x0100-0CCD-CDD0.

```
<PE1> system-view
```

```
[PE1] bpdu-tunnel tunnel-dmac 0100-0ccd-cdd0
```

```

# Create VLAN 2 and assign GigabitEthernet 1/0/1 to VLAN 2.
[PE1] vlan 2
[PE1-vlan2] quit
[PE1] interface gigabitEthernet 1/0/1
[PE1-GigabitEthernet1/0/1] port access vlan 2
# Disable STP on GigabitEthernet 1/0/1, and then enable BPDU tunneling for STP on it.
[PE1-GigabitEthernet1/0/1] undo stp enable
[PE1-GigabitEthernet1/0/1] bpdu-tunnel dot1q stp

```

2. Configure PE 2:

```

# Configure the destination multicast MAC address for BPDUs as 0x0100-0CCD-CDD0.
<PE2> system-view
[PE2] bpdu-tunnel tunnel-dmac 0100-0ccd-cdd0
# Create VLAN 2 and assign GigabitEthernet 1/0/2 to VLAN 2.
[PE2] vlan 2
[PE2-vlan2] quit
[PE2] interface gigabitEthernet 1/0/2
[PE2-GigabitEthernet1/0/2] port access vlan 2
# Disable STP on GigabitEthernet 1/0/2, and then enable BPDU tunneling for STP on it.
[PE2-GigabitEthernet1/0/2] undo stp enable
[PE2-GigabitEthernet1/0/2] bpdu-tunnel dot1q stp

```

BPDU tunneling for PVST configuration example

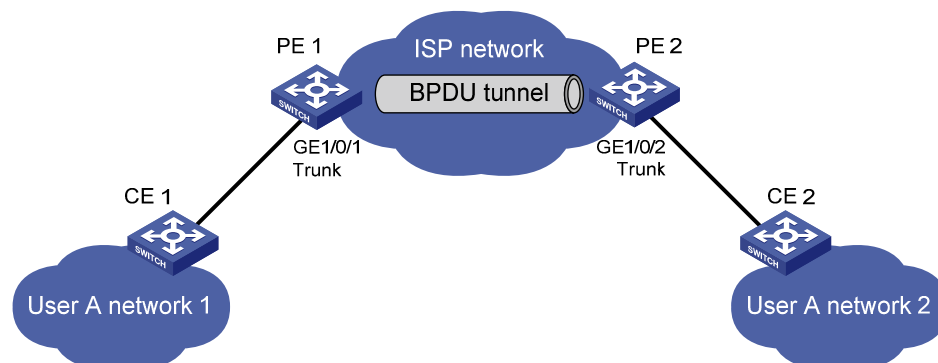
Network requirements

As shown in [Figure 35](#):

- CE 1 and CE 2 are edge devices on the geographically dispersed network of User A. PE 1 and PE 2 are edge devices on the service provider network.
- All ports that connect service provider devices and customer devices and those that interconnect service provider devices are trunk ports and allow packets of any VLAN to pass through.
- PVST is enabled for VLANs 1 through 4094 on User A's network.

After the configuration, CE 1 and CE 2 must implement consistent PVST calculation across the service provider network, and the destination multicast MAC address carried in BPDUs must be 0x0100-0CCD-CDD0.

Figure 35 Network diagram



Configuration procedure

1. Configure PE 1:

Configure the destination multicast MAC address for BPDUs as 0x0100-0CCD-CDD0.

```
<PE1> system-view
```

```
[PE1] bpdu-tunnel tunnel-dmac 0100-0ccd-cdd0
```

Configure GigabitEthernet 1/0/1 as a trunk port and assign it to all VLANs.

```
[PE1] interface gigabitethernet 1/0/1
```

```
[PE1-GigabitEthernet1/0/1] port link-type trunk
```

```
[PE1-GigabitEthernet1/0/1] port trunk permit vlan all
```

Disable STP on GigabitEthernet 1/0/1, and then enable BPDU tunneling for STP and PVST on it.

```
[PE1-GigabitEthernet1/0/1] undo stp enable
```

```
[PE1-GigabitEthernet1/0/1] bpdu-tunnel dot1q stp
```

```
[PE1-GigabitEthernet1/0/1] bpdu-tunnel dot1q pvst
```

2. Configure PE 2:

Configure the destination multicast MAC address for BPDUs as 0x0100-0CCD-CDD0.

```
<PE2> system-view
```

```
[PE2] bpdu-tunnel tunnel-dmac 0100-0ccd-cdd0
```

Configure GigabitEthernet 1/0/2 as a trunk port and assign it to all VLANs.

```
[PE2] interface gigabitethernet 1/0/2
```

```
[PE2-GigabitEthernet1/0/2] port link-type trunk
```

```
[PE2-GigabitEthernet1/0/2] port trunk permit vlan all
```

Disable STP on GigabitEthernet 1/0/2, and then enable BPDU tunneling for STP and PVST on it.

```
[PE2-GigabitEthernet1/0/2] undo stp enable
```

```
[PE2-GigabitEthernet1/0/2] bpdu-tunnel dot1q stp
```

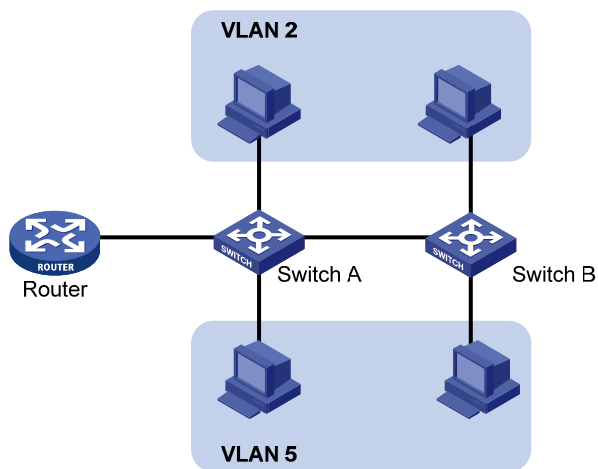
```
[PE2-GigabitEthernet1/0/2] bpdu-tunnel dot1q pvst
```

Configuring VLANs

Overview

Ethernet is a network technology based on the Carrier Sense Multiple Access/Collision Detect (CSMA/CD) mechanism. Because the medium is shared, collisions and excessive broadcasts are common on Ethernet networks. To address the issue, virtual LAN (VLAN) was introduced to break a LAN down into separate VLANs. VLANs are isolated from each other at Layer 2. A VLAN is a bridging domain, and contains all broadcast traffic within it.

Figure 36 A VLAN diagram



A VLAN is logically divided on an organizational basis rather than on a physical basis. For example, using VLAN, all workstations and servers that a particular workgroup uses can be assigned to the same VLAN, regardless of their physical locations.

VLAN technology delivers the following benefits:

1. Confining broadcast traffic within individual VLANs. This reduces bandwidth waste and improves network performance.
2. Improving LAN security. By assigning user groups to different VLANs, you can isolate them at Layer 2. To enable communication between VLANs, routers or Layer 3 switches are required.
3. Creating flexible virtual workgroups. Because users from the same workgroup can be assigned to the same VLAN regardless of their physical locations, network construction and maintenance are much easier and more flexible.

VLAN fundamentals

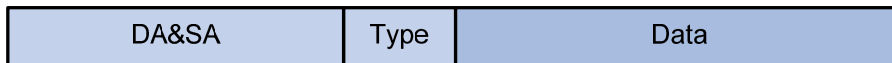
To enable a network device to identify frames of different VLANs, a VLAN tag field is inserted into the data link layer encapsulation.

The format of VLAN-tagged frames is defined in IEEE 802.1Q issued by the Institute of Electrical and Electronics Engineers (IEEE) in 1999.

The Ethernet II encapsulation format is used here. Besides the Ethernet II encapsulation format, Ethernet also supports other encapsulation formats, including 802.2 LLC, 802.2 SNAP, and 802.3 raw. The VLAN tag fields are added to frames encapsulated in these formats for VLAN identification.

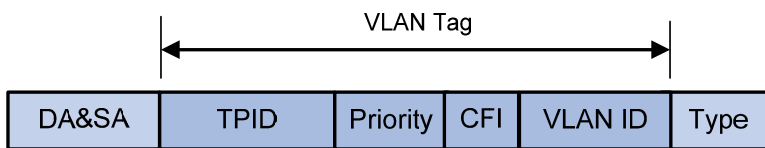
In the header of a traditional Ethernet data frame, the field after the destination MAC address and the source MAC address is the Type field, which indicates the upper layer protocol type, as shown in [Figure 37](#).

Figure 37 Traditional Ethernet frame format



IEEE 802.1Q inserts a four-byte VLAN tag after the DA&SA field, as shown in [Figure 38](#).

Figure 38 Position and format of VLAN tag



The fields of a VLAN tag are tag protocol identifier (TPID), priority, canonical format indicator (CFI), and VLAN ID.

- The 16-bit TPID field with a value of 0x8100 indicates that the frame is VLAN-tagged.
- The 3-bit priority field indicates the 802.1p priority of the frame.
- The 1-bit CFI field specifies whether the MAC addresses are encapsulated in the standard format when packets are transmitted across different media. A value of 0 indicates that MAC addresses are encapsulated in the standard format. A value of 1 indicates that MAC addresses are encapsulated in a non-standard format. The value of the field is 0 by default.
- The 12-bit VLAN ID field identifies the VLAN that the frame belongs to. The VLAN ID range is 0 to 4095. Because 0 and 4095 are reserved, a VLAN ID actually ranges from 1 to 4094.

A network device handles an incoming frame depending on whether the frame is VLAN tagged, and the value of the VLAN tag, if any. For more information, see "[Introduction to port-based VLAN.](#)"

NOTE:

When a frame carrying multiple VLAN tags passes through, the switch processes the frame according to its outer VLAN tag, and transmits the inner tags as payload.

VLAN types

You can implement VLANs based on the following criteria:

- Port
- MAC address
- Protocol
- IP subnet
- Policy
- Other criteria

This chapter covers port-based VLAN, MAC-based VLAN, protocol-based VLAN, and IP subnet-based VLAN. The port-based VLAN implementation is the basis of all other VLAN implementations. To use any other VLAN implementations, you must configure port-based VLAN settings.

You can configure all these types of VLANs on a port at the same time. When the switch is determining which VLAN a packet that passes through the port should be assigned to, it looks up the VLANs in the default order of MAC-based VLAN, IP sub-based VLAN, protocol-based VLAN, and port-based VLAN.

Protocols and standards

IEEE 802.1Q, *IEEE Standards for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks*

Configuring basic VLAN settings

Configuration restrictions and guidelines

- As the default VLAN, VLAN 1 cannot be created or removed.
- You cannot manually create or remove VLANs reserved for special purposes.
- To delete a protocol reserved VLAN, voice VLAN, management VLAN, dynamic VLAN, VLAN with a QoS policy applied, control VLAN for a smart link group, control VLAN for an RRPP domain, remote probe VLAN for remote port mirroring, remove the configuration from the VLAN first, and execute the **undo vlan** command.

Configuration procedure

To configure basic VLAN settings:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create VLANs.	vlan { <i>vlan-id1</i> [to <i>vlan-id2</i>] all }	Optional. Use this command to create VLANs in bulk.
3. Enter VLAN view.	vlan <i>vlan-id</i>	By default, only the default VLAN (VLAN 1) exists in the system. If the specified VLAN does not exist, this command creates the VLAN first.
4. Configure a name for the VLAN.	name <i>text</i>	Optional. By default, the name of a VLAN is its VLAN ID (VLAN 0001 , for example).
5. Configure the description of the VLAN.	description <i>text</i>	Optional. VLAN ID is used by default. (VLAN 0001 , for example).

Configuring basic settings of a VLAN interface

For hosts of different VLANs to communicate, you must use a router or Layer 3 switch to perform Layer 3 forwarding. You use VLAN interfaces to achieve this.

VLAN interfaces are virtual interfaces used for Layer 3 communication between different VLANs. They do not exist as physical entities on devices. For each VLAN, you can create one VLAN interface. You can assign the VLAN interface an IP address and specify it as the gateway of the VLAN to forward traffic destined for an IP subnet different from that of the VLAN.

Configuration procedure

To configure basic settings of a VLAN interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a VLAN interface and enter VLAN interface view.	interface vlan-interface <i>vlan-interface-id</i>	If the VLAN interface already exists, you enter its view directly.
3. Assign an IP address to the VLAN interface.	ip address <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } [<i>sub</i>]	Optional. By default, no IP address is assigned to any VLAN interface.
4. Configure the description of the VLAN interface.	description <i>text</i>	Optional. By default, the description of a VLAN is the VLAN interface name. For example, Vlan-interface1 Interface .
5. Set the MTU for the VLAN interface.	mtu <i>size</i>	Optional. By default, the MTU is 1500 bytes.
6. Restore the default settings for the VLAN interface.	default	Optional.
7. Shut down the VLAN interface.	shutdown	Optional. By default, a VLAN interface is in the up state. The VLAN interface is up if one or more ports in the VLAN is up, and goes down if all ports in the VLAN go down. A VLAN interface shut down with the shutdown command is in the DOWN (Administratively) state until you bring it up, regardless of how the state of the ports in the VLAN changes.

NOTE:

Before you create a VLAN interface for a VLAN, create the VLAN.

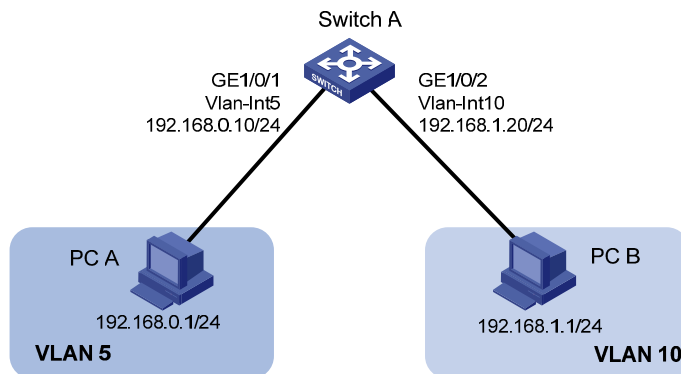
VLAN interface configuration example

Network requirements

As shown in [Figure 39](#), PC A is assigned to VLAN 5. PC B is assigned to VLAN 10. The PCs belong to different IP subnets and cannot communicate with each other.

Configure VLAN interfaces on Switch A and configure the PCs to enable Layer 3 communication between the PCs.

Figure 39 Network diagram



Configuration procedure

1. Configure Switch A:

Create VLAN 5 and assign GigabitEthernet 1/0/1 to it.

```
<SwitchA> system-view
```

```
[SwitchA] vlan 5
```

```
[SwitchA-vlan5] port GigabitEthernet 1/0/1
```

Create VLAN 10 and assign GigabitEthernet 1/0/2 to it.

```
[SwitchA-vlan5] vlan 10
```

```
[SwitchA-vlan10] port GigabitEthernet 1/0/2
```

```
[SwitchA-vlan10] quit
```

Create VLAN-interface 5 and configure its IP address as 192.168.0.10/24.

```
[SwitchA] interface vlan-interface 5
```

```
[SwitchA-Vlan-interface5] ip address 192.168.0.10 24
```

```
[SwitchA-Vlan-interface5] quit
```

Create VLAN-interface 10 and configure its IP address as 192.168.1.20/24.

```
[SwitchA] interface vlan-interface 10
```

```
[SwitchA-Vlan-interface10] ip address 192.168.1.20 24
```

```
[SwitchA-Vlan-interface10] return
```

2. Configure PC A:

Configure the default gateway of the PC as 192.168.0.10.

3. Configure PC B:

Configure the default gateway of the PC as 192.168.1.20.

Verifying the configurations

1. The PCs can ping each other.

2. Display brief information about Layer 3 interfaces on Switch A to verify the configuration.

```
<SwitchA> display ip interface brief
*down: administratively down
(s): spoofing
Interface                Physical Protocol IP Address      Description
Vlan-interface5         up        up        192.168.0.10    Vlan-inte...
Vlan-interface10       up        up        192.168.1.20    Vlan-inte...
```

Configuring port-based VLANs

Introduction to port-based VLAN

Port-based VLANs group VLAN members by port. A port forwards traffic for a VLAN only after it is assigned to the VLAN.

Port link type

You can configure the link type of a port as access, trunk, or hybrid. The link types use the following VLAN tag handling methods:

- An access port belongs to only one VLAN and sends traffic untagged. It is usually used to connect a terminal device unable to identify VLAN tagged-packets or when separating different VLAN members is unnecessary.
- A trunk port can carry multiple VLANs to receive and send traffic for them. Except traffic from the port VLAN ID (PVID), traffic sent through a trunk port will be VLAN tagged. Usually, ports that connect network devices are configured as trunk ports.
- Like a trunk port, a hybrid port can carry multiple VLANs to receive and send traffic for them. Unlike a trunk port, a hybrid port allows traffic of all VLANs to pass through VLAN untagged. You can configure a port connected to a network device or user terminal as a hybrid port.

PVID

By default, VLAN 1 is the PVID for all ports. You can configure the PVID for a port as required.

When you configure the PVID on a port, use the following guidelines:

- An access port can join only one VLAN. The VLAN to which the access port belongs is the PVID of the port. The PVID of the access port changes along with the VLAN to which the port belongs.
- A trunk or hybrid port can join multiple VLANs. You can configure a PVID for the port.
- You can use a nonexistent VLAN as the PVID for a hybrid or trunk port but not for an access port. After you use the **undo vlan** command to remove the VLAN that an access port resides in, the PVID of the port changes to VLAN 1. The removal of the VLAN specified as the PVID of a trunk or hybrid port, however, does not affect the PVID setting on the port.

When you configure a PVID, follow these guidelines:

- Do not set the voice VLAN as the PVID of a port in automatic voice VLAN assignment mode. For information about voice VLAN, see "[Configuring a voice VLAN](#)."
- HP recommends that you set the same PVID ID for local and remote ports.
- Make sure that a port is assigned to its PVID. Otherwise, when the port receives frames tagged with the PVID or untagged frames (including protocol packets such as MSTP BPDUs), the port filters out these frames.

The following table shows how ports of different link types handle frames:

Port type	Actions (in the inbound direction)		Actions (in the outbound direction)
	Untagged frame	Tagged frame	
Access	Tags the frame with the PVID tag.	<ul style="list-style-type: none"> Receives the frame if its VLAN ID is the same as the PVID. Drops the frame if its VLAN ID is different from the PVID. 	Removes the VLAN tag and sends the frame.
Trunk	<p>Checks whether the PVID is permitted on the port:</p> <ul style="list-style-type: none"> If yes, tags the frame with the PVID tag. If not, drops the frame. 	<ul style="list-style-type: none"> Receives the frame if its VLAN is carried on the port. Drops the frame if its VLAN is not carried on the port. 	<ul style="list-style-type: none"> Removes the tag and send the frame if the frame carries the PVID tag and the port belongs to the PVID. Sends the frame without removing the tag if its VLAN is carried on the port but is different from the PVID.
Hybrid			Sends the frame if its VLAN is carried on the port. The frame is sent with the VLAN tag removed or intact depending on your configuration via the port hybrid vlan command. This is true of the PVID.

Assigning an access port to a VLAN

You can assign an access port to a VLAN in VLAN view, interface view (including Layer 2 Ethernet interface view, and Layer 2 aggregate interface view), or port group view.

To assign one or multiple access ports to a VLAN in VLAN view:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	If the specified VLAN does not exist, this command creates the VLAN first.
3. Assign one or a group of access ports to the VLAN.	port <i>interface-list</i>	By default, all ports belong to VLAN 1.

To assign an access port (in interface view) or multiple access ports (in port group view) to a VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i> Enter Layer 2 aggregate interface view: interface bridge-aggregation <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	<p>Use any command.</p> <ul style="list-style-type: none"> The configuration made in Layer 2 Ethernet interface view applies only to the port. The configuration made in port group view applies to all ports in the port group. The configuration made in Layer 2 aggregate interface view applies to the aggregate interface and its aggregation member ports. If the system fails to apply the configuration to the aggregate interface, it stops applying the configuration to aggregation member ports. If the system fails to apply the configuration to an aggregation member port, it skips the port and moves to the next member port.
3. Configure the link type of the ports as access.	port link-type access	Optional. By default, all ports are access ports.
4. Assign the access ports to a VLAN.	port access vlan <i>vlan-id</i>	Optional. By default, all access ports belong to VLAN 1.

NOTE:

- Before you assign an access port to a VLAN, create the VLAN.
- In VLAN view, you can assign only Layer 2 Ethernet interfaces to the VLAN.

Assigning a trunk port to a VLAN

A trunk port can carry multiple VLANs. You can assign it to a VLAN in interface view (including Layer 2 Ethernet interface view, and Layer 2 aggregate interface view) or port group view.

To assign a trunk port to one or multiple VLANs:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i> Enter Layer 2 aggregate interface view: interface bridge-aggregation <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	<p>Use any command.</p> <ul style="list-style-type: none"> The configuration made in Layer 2 Ethernet interface view applies only to the port. The configuration made in port group view applies to all ports in the port group. The configuration made in Layer 2 aggregate interface view applies to the aggregate interface and its aggregation member ports. If the system fails to apply the configuration to the aggregate interface, it stops applying the configuration to aggregation member ports. If the system fails to apply the configuration to an aggregation member port, it skips the port and moves to the next member port.
3. Configure the link type of the ports as trunk.	port link-type trunk	<p>By default, all ports are access ports.</p> <p>To change the link type of a port from trunk to hybrid or vice versa, you must set the link type to access first.</p>
4. Assign the trunk ports to the specified VLANs.	port trunk permit vlan { <i>vlan-id-list</i> all }	By default, a trunk port carries only VLAN 1.
5. Configure the PVID of the trunk ports.	port trunk pvid vlan <i>vlan-id</i>	<p>Optional.</p> <p>By default, the PVID is VLAN 1.</p>

NOTE:

After configuring the PVID for a trunk port, you must use the **port trunk permit vlan** command to configure the trunk port to allow packets from the PVID to pass through, so that the egress port can forward packets from the PVID.

Assigning a hybrid port to a VLAN

A hybrid port can carry multiple VLANs. You can assign it to a VLAN in interface view (including Ethernet interface view, and Layer 2 aggregate interface view) or port group view.

To assign a hybrid port to one or multiple VLANs:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
		Use any command.
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i> Enter Layer 2 aggregate interface view: interface bridge-aggregation <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	<ul style="list-style-type: none"> The configuration made in Ethernet interface view applies only to the port. The configuration made in port group view applies to all ports in the port group. The configuration made in Layer 2 aggregate interface view applies to the aggregate interface and its aggregation member ports. If the system fails to apply the configuration to the aggregate interface, it stops applying the configuration to aggregation member ports. If the system fails to apply the configuration to an aggregation member port, it skips the port and moves to the next member port.
3. Configure the link type of the ports as hybrid.	port link-type hybrid	By default, all ports are access ports. To change the link type of a port from trunk to hybrid or vice versa, you must set the link type to access first.
4. Assign the hybrid ports to the specified VLANs.	port hybrid vlan <i>vlan-id-list</i> { tagged untagged }	By default, a hybrid port allows only packets of VLAN 1 to pass through untagged.
5. Configure the PVID of the hybrid ports.	port hybrid pvid vlan <i>vlan-id</i>	Optional. By default, the PVID is VLAN 1.

NOTE:

- Before you assign a hybrid port to a VLAN, create the VLAN.
- After configuring the PVID for a hybrid port, you must use the **port hybrid vlan** command to configure the hybrid port to allow packets from the PVID to pass through, so that the egress port can forward packets from the PVID.

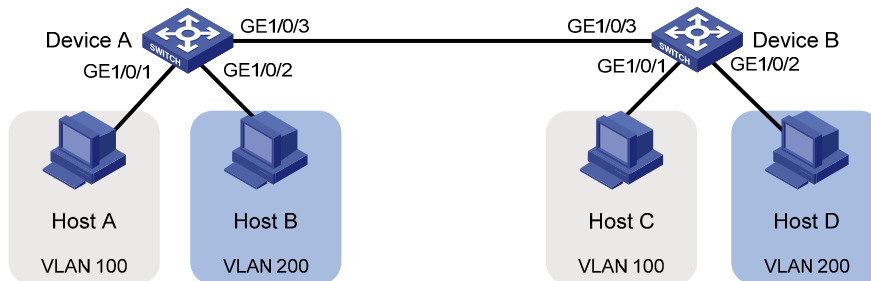
Port-based VLAN configuration example

Network requirements

As shown in [Figure 40](#):

- Host A and Host C belong to Department A, and access the enterprise network through different devices. Host B and Host D belong to Department B. They also access the enterprise network through different devices.
- To ensure communication security and avoid broadcast storms, VLANs are configured in the enterprise network to isolate Layer 2 traffic of different departments. VLAN 100 is assigned to Department A, and VLAN 200 is assigned to Department B.
- Make sure that hosts within the same VLAN can communicate with each other. Host A can communicate with Host C, and Host B can communicate with Host D.

Figure 40 Network diagram



Configuration procedure

1. Configure Device A:

Create VLAN 100, and assign port GigabitEthernet 1/0/1 to VLAN 100.

```
<DeviceA> system-view
[DeviceA] vlan 100
[DeviceA-vlan100] port gigabitethernet 1/0/1
[DeviceA-vlan100] quit
```

Create VLAN 200, and assign port GigabitEthernet 1/0/2 to VLAN 200.

```
[DeviceA] vlan 200
[DeviceA-vlan200] port gigabitethernet 1/0/2
[DeviceA-vlan200] quit
```

Configure port GigabitEthernet 1/0/3 as a trunk port, and assign it to VLANs 100 and 200, to enable GigabitEthernet 1/0/3 to forward traffic of VLANs 100 and 200 to Device B.

```
[DeviceA] interface gigabitethernet 1/0/3
[DeviceA-GigabitEthernet1/0/3] port link-type trunk
[DeviceA-GigabitEthernet1/0/3] port trunk permit vlan 100 200
Please wait... Done.
```

2. Configure Device B as you configure Device A.

3. Configure Host A and Host C to be on the same IP subnet, 192.168.100.0/24, for example. Configure Host B and Host D to be on the same IP subnet, 192.168.200.0/24, for example.

Verifying the configurations

1. Host A and Host C and ping each other successfully, but they both fail to ping Host B. Host B and Host D and ping each other successfully, but they both fail to ping Host A.

2. Determine whether the configuration is successful by displaying relevant VLAN information.

Display information about VLANs 100 and 200 on Device A.

```
[DeviceA-GigabitEthernet1/0/3] display vlan 100
VLAN ID: 100
VLAN Type: static
Route Interface: not configured
Description: VLAN 0100
Name: VLAN 0100
Tagged Ports:
  GigabitEthernet1/0/3
Untagged Ports:
  GigabitEthernet1/0/1
```

```
[DeviceA-GigabitEthernet1/0/3] display vlan 200
VLAN ID: 200
VLAN Type: static
Route Interface: not configured
Description: VLAN 0200
Name: VLAN 0200
Tagged Ports:
  GigabitEthernet1/0/3
Untagged Ports:
  GigabitEthernet1/0/2
```

Configuring MAC-based VLANs

Introduction to MAC-based VLAN

The MAC-based VLAN feature assigns hosts to a VLAN based on their MAC addresses. This feature is usually used in conjunction with security technologies such as 802.1X to provide secure, flexible network access for terminal devices.

Static MAC-based VLAN assignment

Static MAC-based VLAN assignment applies to networks containing a small number of VLAN users. In such a network, you can create a MAC address-to-VLAN map containing multiple MAC address-to-VLAN entries on a port, enable the MAC-based VLAN feature on the port, and assign the port to MAC-based VLANs.

With static MAC-based VLAN assignment configured on a port, the device processes received frames by using the following guidelines:

- When the port receives an untagged frame, the device looks up the MAC address-to-VLAN map based on the source MAC address of the frame for a match.
 - a. The device first performs a fuzzy match. In the fuzzy match, the device searches the MAC address-to-VLAN entries whose masks are not all-Fs and performs a logical AND operation on the source MAC address and each mask. If the result of an AND operation matches the corresponding MAC address, the device tags the frame with the corresponding VLAN ID. Only the 5500 EI Switch Series supports fuzzy match.
 - b. If the fuzzy match fails, the device performs an exact match. In the exact match, the device searches the MAC address-to-VLAN entries whose masks are all-Fs. If the MAC address of a MAC address-to-VLAN entry matches the source MAC address of the untagged frame, the device tags the frame with the corresponding VLAN ID.
 - c. If no match is found, the device assigns a VLAN to the frame by using other criteria, such as IP subnet or protocol, and forwards the frame.
 - d. If no VLAN is available, the device tags the frame with the PVID of the receiving port and forwards the frame.
- When the port receives a tagged frame, the port forwards the frame if the VLAN ID of the frame is permitted by the port, or otherwise drops the frame.

Dynamic MAC-based VLAN assignment

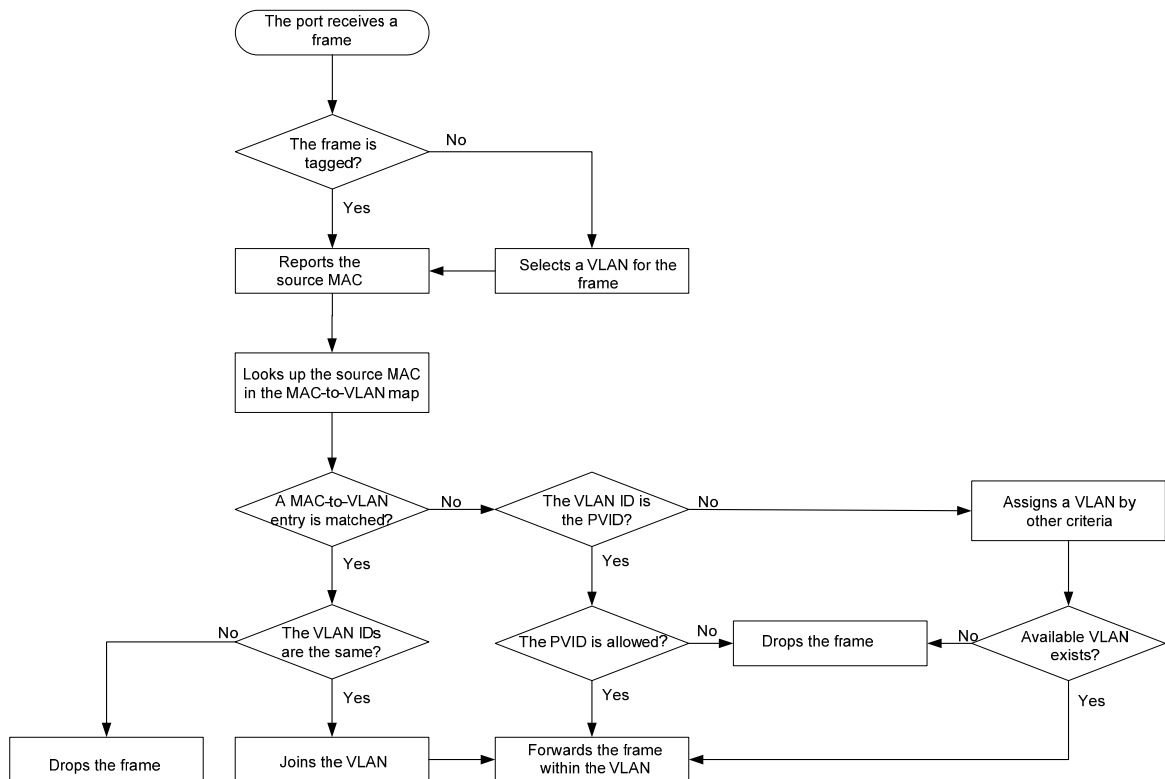
When you cannot determine the target MAC-based VLANs of a port, you can use dynamic MAC-based VLAN assignment on the port. To do that, you can create a MAC address-to-VLAN map containing

multiple MAC address-to-VLAN entries, and enable the MAC-based VLAN feature and dynamic MAC-based VLAN assignment on the port.

Dynamic MAC-based VLAN assignment uses the following workflows.

1. When the port receives a frame, the port first determines whether the frame is tagged.
 - If yes, the port reports the source MAC address of the frame.
 - If not, the port selects a VLAN for the frame by tagging the untagged frame with the PVID tag and obtaining the tag, and then reports the source MAC address of the frame.
2. After reporting the source MAC address of the frame, the port looks up the source MAC address in the MAC-to-VLAN map, and processes the frame as follows:
 - If the source MAC address of the frame matches a MAC address-to-VLAN entry configured on the port, the port checks whether the VLAN ID of the frame is the same as the VLAN in the MAC-to-VLAN entry.
 - a. If yes, the port dynamically joins the VLAN and forwards the frame.
 - b. If not, the port drops the frame.
 - If the source MAC address of the frame matches no MAC-to-VLAN entry, the port processes the frame depending on whether the VLAN ID of the frame is the PVID.
 - c. If yes, the port determines whether it allows PVID: if yes, the port tags the frame with the PVID and forwards the frame; if not, the port drops the frame.
 - d. If not, the port assigns a VLAN to the frame by using other criteria, such as IP subnet or protocol, and forwards the frame. If no VLAN is available, the port drops the frame.

Figure 41 Flowchart for processing a frame in dynamic MAC-based VLAN assignment



When you configure dynamic MAC-based VLAN assignment, follow these guidelines:

- When a port is assigned to the corresponding VLAN in a MAC address-to-VLAN entry, but has not been assigned to the VLAN by using the **port hybrid vlan** command, the port sends packets from the VLAN with VLAN tags removed.
- If you configure both static and dynamic MAC-based VLAN assignment on the same port, dynamic MAC-based VLAN assignment applies.

Dynamic MAC-based VLAN

You can use dynamic MAC-based VLAN with access authentication (such as 802.1X authentication based on MAC addresses) to implement secure, flexible terminal access. After configuring dynamic MAC-based VLAN on the device, you must configure the username-to-VLAN entries on the access authentication server.

When a user passes authentication of the access authentication server, the device obtains VLAN information from the server, generates a MAC address-to-VLAN entry by using the source MAC address of the user packet and the VLAN information, and assigns the port to the MAC-based VLAN. When the user goes offline, the device automatically deletes the MAC address-to-VLAN entry, and removes the port from the MAC-based VLAN. For more information about 802.1X, MAC, and portal authentication, see *Security Configuration Guide*.

Configuration restrictions and guidelines

When you configure a MAC-based VLAN, follow these guidelines:

- MAC-based VLANs are available only on hybrid ports.
- You cannot configure super VLANs in the MAC address-to-VLAN entries.
- With dynamic MAC-based VLAN assignment enabled, packets are delivered to the CPU for processing. The packet processing mode has the highest priority and overrides the configuration of MAC learning limit and disabling of MAC address learning. When dynamic MAC-based VLAN assignment is enabled, do not configure the MAC learning limit or disable MAC address learning.
- Do not use dynamic MAC-based VLAN assignment together with 802.X and MAC authentication.
- In dynamic MAC-based VLAN assignment, the port that receives a packet with an unknown source MAC address can be successfully assigned to the matched VLAN only when the matched VLAN is a static VLAN.
- The MAC-based VLAN feature is mainly configured on the downlink ports of the user access devices. Do not enable this function together with link aggregation.
- With MSTP enabled, if a port is blocked in the MST instance (MSTI) of the target MAC-based VLAN, the port drops the received packets, instead of delivering them to the CPU. As a result, the receiving port will not be dynamically assigned to the corresponding VLAN. Do not configure dynamic MAC-based VLAN assignment together with MSTP, because the former is mainly configured on the access side.
- When you configure MAC-to-VLAN entries, if you specify the 802.1p priority for the VLAN of a MAC address, you must configure the **qos trust dot1p** command on the corresponding port, so that the port trusts the 802.1p priority of incoming packets and your configuration takes effect. For more information about the **qos trust dot1p** command, see *ACL and QoS Command Reference*.

Configuration procedure

To configure static MAC-based VLAN assignment:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Associate a specific MAC address with a VLAN.	mac-vlan mac-address <i>mac-address</i> [mask <i>mac-mask</i>] vlan <i>vlan-id</i> [priority <i>priority</i>]	The mask keyword is available on only the 5500 EI Switch Series.
3. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	<p>Use either command.</p> <ul style="list-style-type: none"> The configuration made in Ethernet interface view applies only to the port. The configuration made in port group view applies to all ports in the port group.
4. Configure the link type of the ports as hybrid.	port link-type hybrid	By default, all ports are access ports.
5. Configure the hybrid ports to permit packets from specific MAC-based VLANs to pass through.	port hybrid vlan <i>vlan-id-list</i> { tagged untagged }	By default, a hybrid port only permits the packets from VLAN 1 to pass through.
6. Enable the MAC-based VLAN feature.	mac-vlan enable	Disabled by default.
7. Configure VLAN matching precedence.	vlan precedence { mac-vlan ip-subnet-vlan }	Optional. By default, VLANs are preferably matched based on MAC addresses.

To configure dynamic MAC-based VLAN assignment:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Associate MAC addresses with a VLAN.	mac-vlan mac-address <i>mac-address</i> vlan <i>vlan-id</i> [priority <i>priority</i>]	N/A
3. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Configure the link type of the port as hybrid.	port link-type hybrid	By default, all ports are access ports.
5. Enable the MAC-based VLAN feature.	mac-vlan enable	Disabled by default.

Step	Command	Remarks
6. Enable dynamic MAC-based VLAN assignment.	mac-vlan trigger enable	By default, dynamic MAC-based VLAN assignment is disabled. When you use the mac-vlan trigger enable command to enable dynamic MAC-based VLAN assignment, HP recommends that you configure the vlan precedence mac-vlan command, so that VLANs are assigned based on single MAC addresses preferentially. When dynamic MAC-based VLAN assignment is enabled, HP does not recommend configuring the vlan precedence ip-subnet-vlan command, which will make the system assign VLANs based on IP subnets, because the configuration does not take effect.
7. Configure VLAN matching precedence.	vlan precedence mac-vlan	Optional. By default, VLANs are preferentially matched based on MAC addresses.
8. Disable the PVID of the port from forwarding packets with unknown source MAC addresses that do not match any MAC address-to-VLAN entry.	port pvid disable	Optional. By default, when a port receives a packet with an unknown source MAC address that does not match to any MAC address-to-VLAN entry, it forwards the packet in its PVID.

To configure dynamic MAC-based VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view: interface <i>interface-type interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	<p>Use either command.</p> <ul style="list-style-type: none"> The configuration made in Ethernet interface view applies only to the port. The configuration made in port group view applies to all ports in the port group.
3. Configure the link type of the ports as hybrid.	port link-type hybrid	By default, all ports are access ports.
4. Configure the hybrid ports to permit packets from specific MAC-based VLANs to pass through.	port hybrid vlan <i>vlan-id-list</i> { tagged untagged }	By default, a hybrid port only permits the packets of VLAN 1 to pass through.
5. Enable the MAC-based VLAN feature.	mac-vlan enable	Disabled by default.
6. Configure 802.1X/MAC/port authentication or any combination.	For more information, see <i>Security Command Reference</i> .	N/A

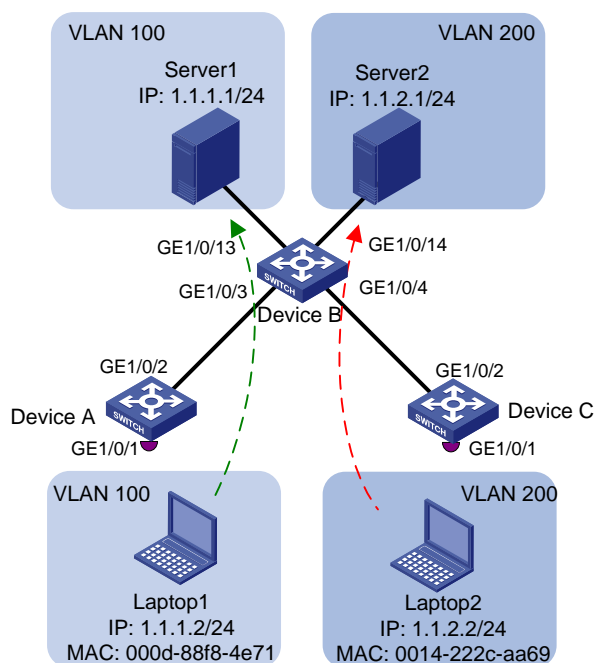
MAC-based VLAN configuration example

Network requirements

As shown in Figure 42:

- GigabitEthernet 1/0/1 of Device A and Device C are each connected to a meeting room. Laptop 1 and Laptop 2 are used for meetings and might be used in either of the two meeting rooms.
- Different departments own Laptop 1 and Laptop 2. The two departments use VLAN 100 and VLAN 200, respectively. Each laptop must be able to access only its own department server, no matter which meeting room it is used in.
- The MAC address of Laptop 1 is 000D-88F8-4E71, and that of Laptop 2 is 0014-222C-AA69.

Figure 42 Network diagram



Configuration consideration

- Create VLANs 100 and 200.
- Configure the uplink ports of Device A and Device C as trunk ports, and assign them to VLANs 100 and 200.
- Configure the downlink ports of Device B as trunk ports, and assign them to VLANs 100 and 200. Assign the uplink ports of Device B to VLANs 100 and 200.
- Associate the MAC address of Laptop 1 with VLAN 100, and associate the MAC address of Laptop 2 with VLAN 200.

Configuration procedure

1. Configure Device A:
Create VLANs 100 and 200.

```
<DeviceA> system-view  
[DeviceA] vlan 100
```

```
[DeviceA-vlan100] quit
[DeviceA] vlan 200
[DeviceA-vlan200] quit

# Associate the MAC address of Laptop 1 with VLAN 100, and associate the MAC address of
Laptop 2 with VLAN 200.
```

```
[DeviceA] mac-vlan mac-address 000d-88f8-4e71 vlan 100
```

```
[DeviceA] mac-vlan mac-address 0014-222c-aa69 vlan 200
```

```
# Configure Laptop 1 and Laptop 2 to access the network through GigabitEthernet 1/0/1.
Configure GigabitEthernet 1/0/1 as a hybrid port that sends packets of VLANs 100 and 200
untagged, and enable the MAC-based VLAN feature on it.
```

```
[DeviceA] interface gigabitethernet 1/0/1
```

```
[DeviceA-GigabitEthernet1/0/1] port link-type hybrid
```

```
[DeviceA-GigabitEthernet1/0/1] port hybrid vlan 100 200 untagged
```

```
Please wait... Done.
```

```
[DeviceA-GigabitEthernet1/0/1] mac-vlan enable
```

```
[DeviceA-GigabitEthernet1/0/1] quit
```

```
# To enable the laptops to access Server 1 and Server 2, configure the uplink port GigabitEthernet
1/0/2 as a trunk port, and assign it to VLANs 100 and 200.
```

```
[DeviceA] interface gigabitethernet 1/0/2
```

```
[DeviceA-GigabitEthernet1/0/2] port link-type trunk
```

```
[DeviceA-GigabitEthernet1/0/2] port trunk permit vlan 100 200
```

```
[DeviceA-GigabitEthernet1/0/2] quit
```

2. Configure Device B:

```
# Create VLANs 100 and 200. Assign GigabitEthernet 1/0/13 to VLAN 100, and assign
GigabitEthernet 1/0/14 to VLAN 200.
```

```
<DeviceB> system-view
```

```
[DeviceB] vlan 100
```

```
[DeviceB-vlan100] port gigabitethernet 1/0/13
```

```
[DeviceB-vlan100] quit
```

```
[DeviceB] vlan 200
```

```
[DeviceB-vlan200] port gigabitethernet 1/0/14
```

```
[DeviceB-vlan200] quit
```

```
# Configure GigabitEthernet 1/0/3 and GigabitEthernet 1/0/4 as trunk ports, and assign them
to VLANs 100 and 200.
```

```
[DeviceB] interface gigabitethernet 1/0/3
```

```
[DeviceB-GigabitEthernet1/0/3] port link-type trunk
```

```
[DeviceB-GigabitEthernet1/0/3] port trunk permit vlan 100 200
```

```
[DeviceB-GigabitEthernet1/0/3] quit
```

```
[DeviceB] interface gigabitethernet 1/0/4
```

```
[DeviceB-GigabitEthernet1/0/4] port link-type trunk
```

```
[DeviceB-GigabitEthernet1/0/4] port trunk permit vlan 100 200
```

```
[DeviceB-GigabitEthernet1/0/4] quit
```

3. Configure Device C:

```
Configure Device C as you configure Device A.
```

Verifying the configurations

1. Laptop 1 can access Server 1 only, and Laptop 2 can access Server 2 only.

2. On Device A and Device C, you can see that VLAN 100 is associated with the MAC address of Laptop 1, and VLAN 200 is associated with the MAC address of Laptop 2.

```
[DeviceA] display mac-vlan all
The following MAC VLAN addresses exist:
S:Static D:Dynamic
MAC ADDR          MASK                VLAN ID  Prio  STATE
-----
000d-88f8-4e71    ffff-ffff-ffff    100      0     S
0014-222c-aa69    ffff-ffff-ffff    200      0     S

Total MAC VLAN address count:2
```

Configuration guidelines

1. MAC-based VLAN can be configured only on hybrid ports.
2. MAC-based VLAN is usually configured on the downlink ports of access layer devices, and cannot be configured together with the link aggregation function.

Configuring protocol-based VLANs

Introduction to protocol-based VLAN

You use the protocol-based VLAN feature to assign packets to VLANs by their application type.

The protocol-based VLAN feature assigns inbound packets to different VLANs based on their protocol type and encapsulation format. The protocols available for VLAN assignment include IP, IPX, and AppleTalk (AT), and the encapsulation formats include Ethernet II, 802.3 raw, 802.2 LLC, and 802.2 SNAP.

A protocol template defines a protocol type and an encapsulation format. A protocol-based VLAN ID and a protocol index, combined, can uniquely identify a protocol template. You can assign multiple protocol templates to a protocol-based VLAN.

Protocol-based VLAN assignment is available only on hybrid ports, and a protocol template applies only to untagged packets.

When an untagged packet arrives, a protocol-based VLAN assignment enabled hybrid port processes the packet by using the following workflow:

- If the protocol type and encapsulation format in the packet matches a protocol template, the packet is tagged with the VLAN tag specific to the protocol template.
- If no protocol template is matched, the packet is tagged with the PVID of the port.

The port processes a tagged packet as it processes tagged packets of a port-based VLAN.

- If the port is in the same VLAN as the packet, it forwards the packet.
- If not, the port drops the packet.

Configuration restrictions and guidelines

A protocol-based VLAN processes only untagged inbound packets, whereas the voice VLAN in automatic mode processes only tagged voice traffic. Do not configure a VLAN as both a protocol-based VLAN and a voice VLAN. For more information, see "[Configuring a voice VLAN.](#)"

Configuration procedure

To configure a protocol-based VLAN:

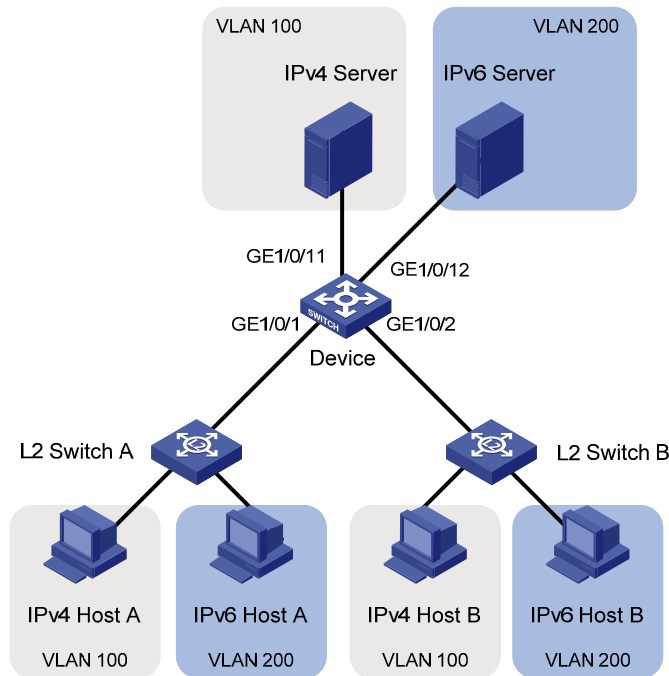
Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	If the specified VLAN does not exist, this command creates the VLAN first.
3. Create a protocol template for the VLAN.	protocol-vlan [<i>protocol-index</i>] { at ipv4 ipv6 ipx { ethernetii llc raw snap } mode { ethernetii etertype <i>etype-id</i> llc { dsap <i>dsap-id</i> ssap <i>ssap-id</i> } snap etertype <i>etype-id</i> }	Not configured by default.
4. Exit VLAN view.	quit	N/A
5. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view: interface <i>interface-type interface-number</i> Enter Layer 2 aggregate interface view: interface bridge-aggregation <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	<p>Use any command.</p> <ul style="list-style-type: none"> The configuration made in Ethernet interface view applies only to the port. The configuration made in port group view applies to all ports in the port group. The configuration made in Layer 2 aggregate interface view applies to the aggregate interface and its aggregation member ports. If the system fails to apply the configuration to the aggregate interface, it stops applying the configuration to aggregation member ports. If the system fails to apply the configuration to an aggregation member port, it skips the port and moves to the next member port.
6. Configure the port link type as hybrid.	port link-type hybrid	By default, all ports are access ports.
7. Assign the hybrid port to the specified protocol-based VLANs.	port hybrid vlan <i>vlan-id-list</i> { tagged untagged }	By default, a hybrid port is only in VLAN 1.
8. Assign the protocol template you have created to the hybrid port.	port hybrid protocol-vlan <i>vlan-id</i> { <i>protocol-index</i> [to <i>protocol-end</i>] all }	N/A

Protocol-based VLAN configuration example

Network requirements

In a lab environment, as shown in [Figure 43](#), most hosts run the IPv4 protocol, and the rest of the hosts run the IPv6 protocol for teaching purposes. To avoid interference, isolate IPv4 traffic and IPv6 traffic at Layer 2.

Figure 43 Network diagram



Configuration consideration

Create VLANs 100 and 200. Associate VLAN 100 with IPv4, and associate VLAN 200 with IPv6. Configure protocol-based VLANs to isolate IPv4 traffic and IPv6 traffic at Layer 2.

Configuration procedure

1. Configure Device:

Create VLAN 100, and assign port GigabitEthernet 1/0/11 to VLAN 100.

```
<Device> system-view
[Device] vlan 100
[Device-vlan100] description protocol VLAN for IPv4
[Device-vlan100] port gigabitethernet 1/0/11
[Device-vlan100] quit
```

Create VLAN 200, and assign port GigabitEthernet 1/0/12 to VLAN 200.

```
[Device] vlan 200
[Device-vlan200] description protocol VLAN for IPv6
[Device-vlan200] port gigabitethernet 1/0/12
```

Create an IPv6 protocol template in the view of VLAN 200, and create an IPv4 protocol template in the view of VLAN 100.

```
[Device-vlan200] protocol-vlan 1 ipv6
[Device-vlan200] quit
[Device] vlan 100
[Device-vlan100] protocol-vlan 1 ipv4
[Device-vlan100] quit
```

Configure port GigabitEthernet 1/0/1 as a hybrid port that forwards packets of VLANs 100 and 200 untagged.

```
[Device] interface gigabitethernet 1/0/1
[Device-GigabitEthernet1/0/1] port link-type hybrid
```

```
[Device-GigabitEthernet1/0/1] port hybrid vlan 100 200 untagged
Please wait... Done.
```

Associate port GigabitEthernet 1/0/1 with the IPv4 protocol template of VLAN 100 and the IPv6 protocol template of VLAN 200.

```
[Device-GigabitEthernet1/0/1] port hybrid protocol-vlan vlan 100 1
[Device-GigabitEthernet1/0/1] port hybrid protocol-vlan vlan 200 1
[Device-GigabitEthernet1/0/1] quit
```

Configure GigabitEthernet 1/0/2 as a hybrid port that forwards packets of VLANs 100 and 200 untagged, and associate GigabitEthernet 1/0/2 with the IPv4 protocol template of VLAN 100 and the IPv6 protocol template of VLAN 200.

```
[Device] interface gigabitethernet 1/0/2
[Device-GigabitEthernet1/0/2] port link-type hybrid
[Device-GigabitEthernet1/0/2] port hybrid vlan 100 200 untagged
Please wait... Done.
[Device-GigabitEthernet1/0/2] port hybrid protocol-vlan vlan 100 1
[Device-GigabitEthernet1/0/2] port hybrid protocol-vlan vlan 200 1
```

2. Keep the default settings of L2 Switch A and L2 Switch B.
3. Configure IPv4 Host A, IPv4 Host B, and IPv4 Server to be on the same IP subnet (192.168.100.0/24, for example), and configure IPv6 Host A, IPv6 Host B, and IPv6 Server to be on the same IP subnet (2001::1/64, for example).

Verifying the configurations

1. The hosts and the server in VLAN 100 can ping one another successfully. The hosts and the server in VLAN 200 can ping one another successfully. The hosts or server in VLAN 100 cannot ping the hosts and server in VLAN 200, and vice versa.
2. Display protocol-based VLAN information on Device to determine whether the configurations have become valid.

Display protocol-based VLAN configuration on Device.

```
[Device-GigabitEthernet1/0/2] display protocol-vlan vlan all
VLAN ID:100
  Protocol Index      Protocol Type
=====
  1                   ipv4
VLAN ID:200
  Protocol Index      Protocol Type
=====
  1                   ipv6
```

Display protocol-based VLAN information on the ports of Device.

```
[Device-GigabitEthernet1/0/2] display protocol-vlan interface all
Interface: GigabitEthernet 1/0/1
  VLAN ID  Protocol Index      Protocol Type
=====
  100      1                   ipv4
  200      1                   ipv6
Interface: GigabitEthernet 1/0/2
  VLAN ID  Protocol Index      Protocol Type
=====
  100      1                   ipv4
```

Configuration guidelines

Protocol-based VLAN configuration applies only to hybrid ports.

Configuring IP subnet-based VLANs

In this approach, packets are assigned to VLANs based on their source IP addresses and subnet masks. A port configured with IP subnet-based VLANs assigns a received untagged packet to a VLAN based on the source address of the packet.

This feature is used to assign packets from the specified IP subnet or IP address to a specific VLAN.

Configuration procedure

⚠ IMPORTANT:

This feature is applicable only on hybrid ports.

To configure an IP subnet-based VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
3. Associate an IP subnet with the VLAN.	ip-subnet-vlan [<i>ip-subnet-index</i>] ip <i>ip-address</i> [<i>mask</i>]	The IP subnet or IP address to be associated with a VLAN cannot be a multicast subnet or a multicast address.
4. Return to system view.	quit	N/A
5. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i> Enter Layer 2 aggregate interface view: interface bridge-aggregation <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	<p>Use any command.</p> <ul style="list-style-type: none"> The configuration made in Ethernet interface view applies only to the port. The configuration made in port group view applies to all ports in the port group. The configuration made in Layer 2 aggregate interface view applies to the aggregate interface and its aggregation member ports. If the system fails to apply the configuration to the aggregate interface, it stops applying the configuration to aggregation member ports. If the system fails to apply the configuration to an aggregation member port, it skips the port and moves to the next member port.
6. Configure port link type as hybrid.	port link-type hybrid	By default, all ports are access ports.

Step	Command	Remarks	
7.	Configure the hybrid ports to permit the specified IP subnet-based VLANs to pass through.	port hybrid vlan <i>vlan-id-list</i> { tagged untagged }	By default, a hybrid port allows only packets from VLAN 1 to pass through untagged.
8.	Associate the hybrid ports with the specified IP subnet-based VLAN.	port hybrid ip-subnet-vlan <i>vlan</i> <i>vlan-id</i>	Not configured by default.

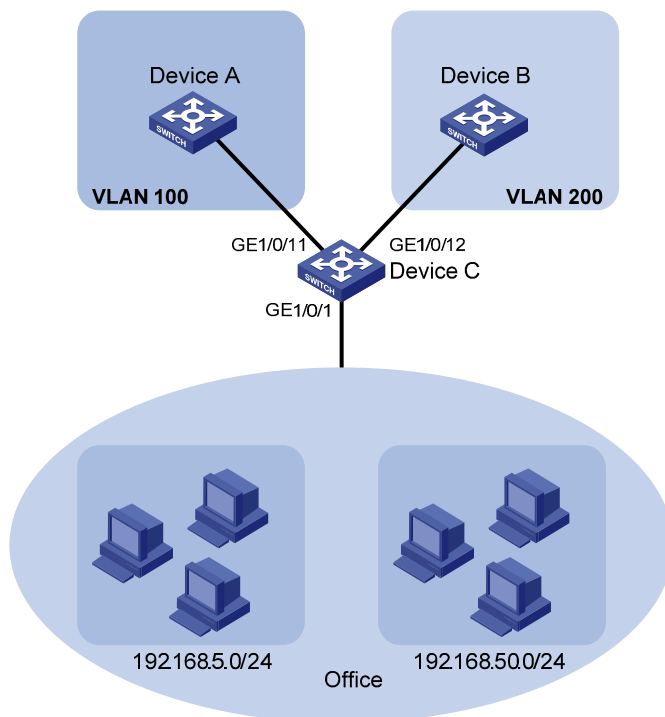
IP subnet-based VLAN configuration example

Network requirements

As shown in Figure 44, the hosts in the office belong to different IP subnets 192.168.5.0/24 and 192.168.50.0/24.

Configure Device C to transmit packets over separate VLANs based on their source IP addresses.

Figure 44 Network diagram



Configuration consideration

- Create VLANs 100 and 200.
- Associate IP subnets with the VLANs.
- Assign ports to the VLANs.

Configuration procedure

Associate IP subnet 192.168.5.0/24 with VLAN 100.

```
<DeviceC> system-view
```

```

[DeviceC] vlan 100
[DeviceC-vlan100] ip-subnet-vlan ip 192.168.5.0 255.255.255.0
[DeviceC-vlan100] quit

# Associate IP subnet 192.168.50.0/24 with VLAN 200.
[DeviceC] vlan 200
[DeviceC-vlan200] ip-subnet-vlan ip 192.168.50.0 255.255.255.0
[DeviceC-vlan200] quit

# Configure interface GigabitEthernet 1/0/11 to permit packets of VLAN 100 to pass through.
[DeviceC] interface GigabitEthernet 1/0/11
[DeviceC-GigabitEthernet1/0/11] port link-type hybrid
[DeviceC-GigabitEthernet1/0/11] port hybrid vlan 100 tagged
Please wait... Done.
[DeviceC-GigabitEthernet1/0/11] quit

# Configure interface GigabitEthernet 1/0/12 to permit packets of VLAN 200 to pass through.
[DeviceC] interface GigabitEthernet 1/0/12
[DeviceC-GigabitEthernet1/0/12] port link-type hybrid
[DeviceC-GigabitEthernet1/0/12] port hybrid vlan 200 tagged
Please wait... Done.
[DeviceC-GigabitEthernet1/0/12] quit

# Associate interface GigabitEthernet 1/0/1 with IP subnet-based VLANs 100 and 200.
[DeviceC] interface GigabitEthernet 1/0/1
[DeviceC-GigabitEthernet1/0/1] port link-type hybrid
[DeviceC-GigabitEthernet1/0/1] port hybrid vlan 100 200 untagged
Please wait... Done.
[DeviceC-GigabitEthernet1/0/1] port hybrid ip-subnet-vlan vlan 100
[DeviceC-GigabitEthernet1/0/1] port hybrid ip-subnet-vlan vlan 200
[DeviceC-GigabitEthernet1/0/1] return

```

Verifying the configurations

```

# Display the IP subnet information for all VLANs.
<Device C> display ip-subnet-vlan vlan all
VLAN ID: 100
Subnet Index      IP Address      Subnet Mask
=====
      0          192.168.5.0    255.255.255.0
VLAN ID: 200
Subnet Index      IP Address      Subnet Mask
=====
      0          192.168.50.0   255.255.255.0

# Display the IP subnet-based VLAN information on GigabitEthernet 1/0/1.
<DeviceC> display ip-subnet-vlan interface GigabitEthernet 1/0/1
Interface: GigabitEthernet1/0/1
VLAN ID  Subnet-Index  IP ADDRESS      NET MASK
=====
    100         0          192.168.5.0    255.255.255.0
    200         0          192.168.50.0   255.255.255.0

```

Configuration guidelines

The IP subnet-based VLAN configurations are only effective on hybrid ports.

Displaying and maintaining VLAN

Task	Command	Remarks
Display VLAN information.	display vlan [<i>vlan-id1</i> [<i>to</i> <i>vlan-id2</i>] all dynamic reserved static] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display VLAN interface information.	display interface [<i>vlan-interface</i>] [brief [down]] [{ begin exclude include } <i>regular-expression</i>] display interface <i>vlan-interface</i> <i>vlan-interface-id</i> [brief] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display hybrid ports or trunk ports on the device.	display port { hybrid trunk } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display MAC address-to-VLAN entries.	display mac-vlan { all dynamic mac-address <i>mac-address</i> [mask <i>mac-mask</i>] static vlan <i>vlan-id</i> } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display all interfaces with MAC-based VLAN enabled.	display mac-vlan interface [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display protocol information and protocol indexes of the specified VLANs.	display protocol-vlan <i>vlan</i> { <i>vlan-id</i> [<i>to</i> <i>vlan-id</i>] all } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display protocol-based VLAN information on specified interfaces.	display protocol-vlan interface { <i>interface-type</i> <i>interface-number</i> [<i>to</i> <i>interface-type</i> <i>interface-number</i>] all } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IP subnet-based VLAN information and IP subnet indexes of specified VLANs.	display ip-subnet-vlan <i>vlan</i> { <i>vlan-id</i> [<i>to</i> <i>vlan-id</i>] all } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the IP subnet-based VLAN information and IP subnet indexes of specified ports.	display ip-subnet-vlan interface { <i>interface-list</i> all } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear statistics on a port.	reset counters interface <i>vlan-interface</i> [<i>vlan-interface-id</i>]	Available in user view

Configuring a super VLAN (available only on the 5500 EI)

Overview

Super VLAN, also called VLAN aggregation, was introduced to save IP address space.

A super VLAN is associated with multiple sub-VLANs. You can create a VLAN interface for a super VLAN and assign an IP address for the VLAN interface. However, you cannot create a VLAN interface for a sub-VLAN. You can assign a physical port to a sub-VLAN, but not to a super VLAN. All ports of a sub-VLAN use the VLAN interface IP address of the associated super VLAN. Packets cannot be forwarded between sub-VLANs at Layer 2.

To enable Layer 3 communication between sub-VLANs, create a super VLAN and the VLAN interface, and enable local proxy Address Resolution Protocol (ARP) or local proxy Neighbor Discovery (ND) on the VLAN interface depending on the VLAN interface IP address type (IPv4 or IPv6) as follows:

- In an IPv4 network, enable local proxy ARP on the VLAN interface, so that the super VLAN can forward and process ARP requests and replies.
- In an IPv6 network, enable local proxy ND on the VLAN interface, so that the super VLAN can forward and process the Neighbor Solicitation (NS) messages and Neighbor Advertisement (NA) messages.

Configuration procedure

To configure a super VLAN, complete the following tasks:

1. Configure sub-VLANs.
2. Configure a super VLAN, and associate the super VLAN with the sub-VLANs configured earlier.
3. Configure a VLAN interface for the super VLAN. The VLAN interface enables communication among hosts and sub-VLANs.

Configuring sub-VLANs

To configure a sub-VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a sub-VLAN and enter VLAN view.	vlan <i>vlan-id</i>	If the specified VLAN already exists, this command enters VLAN view only.

NOTE:

To configure more sub-VLANs, repeat these steps.

Configuring a super VLAN

To configure a super VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	If the specified VLAN does not exist, this command creates the VLAN first, and then enters VLAN view.
3. Configure the VLAN as a super VLAN.	supervlan	Not configured by default.
4. Associate the super VLAN with the specified sub-VLANs.	subvlan <i>vlan-list</i>	VLANs specified by <i>vlan-list</i> must be the sub-VLANs configured earlier.

NOTE:

- When dynamic MAC-based VLAN assignment is enabled on a port, you cannot configure VLANs of the MAC address-to-VLAN entries as super VLANs.
- When a VLAN is configured as an isolate-user-VLAN or associated with secondary VLANs, you cannot configure the VLAN as a super VLAN.

Configure a VLAN interface for the super VLAN

When you a VLAN interface for the super VLAN, follow these guidelines:

- Configure the IP address of the VLAN interface with that of the corresponding super VLAN.
- For more information about local proxy ARP and proxy ND functions, see *Layer 3—IP Services Configuration Guide*. For more information about **local-proxy-arp enable** and **local-proxy-nd enable** commands, see *Layer 3—IP Services Command Reference*.
- You cannot configure a super VLAN as the guest VLAN for a port, and vice versa. For more information about guest VLANs, see *Security Configuration Guide*.
- You can configure Layer 2 multicast for a super VLAN, but the configuration is ineffective.
- You can configure DHCP, Layer 3 multicast, and dynamic routing for the VLAN interface of a super VLAN. However, only DHCP takes effect.
- HP does not recommend configuring VRRP for the VLAN interface of a super VLAN, because it affects network performance. For more information about VRRP, see *High Availability Configuration Guide*.

To configure a VLAN interface for the super VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a VLAN interface, and enter VLAN interface view.	interface vlan-interface <i>vlan-interface-id</i>	The value of <i>vlan-interface-id</i> must be the ID of the super VLAN.
3. Configure the IP address of the VLAN interface.	<ul style="list-style-type: none"> • ip address <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } [sub] • ipv6 address { <i>ipv6-address</i> { <i>prefix-length</i> link-local } <i>ipv6-address/prefix-length</i> [anycast eui-64] auto [link-local] } 	Use either command. By default, the IP address of a VLAN interface is not configured.

Step	Command	Remarks
4.	Enable local proxy ARP. local-proxy-arp enable	Use either command.
5.	Enable local proxy ND. local-proxy-nd enable	Disabled by default.

Displaying and maintaining super VLAN

Task	Command	Remarks
Display the mapping between a super VLAN and its sub-VLANs.	display supervlan [<i>supervlan-id</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

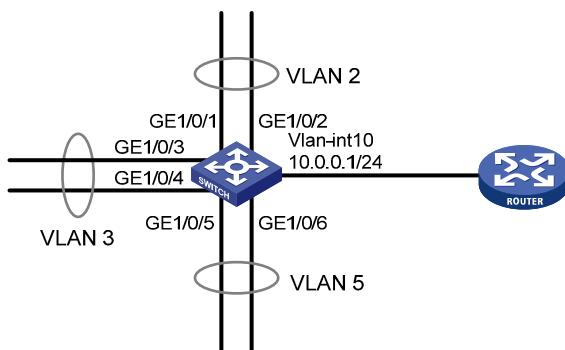
Super VLAN configuration example

Network requirements

As shown in Figure 45:

- Create super VLAN 10, and configure its VLAN interface IP address as 10.0.0.1/24.
- Create the sub-VLANs VLAN 2, VLAN 3, and VLAN 5.
- Assign GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 to VLAN 2, GigabitEthernet 1/0/3 and GigabitEthernet 1/0/4 to VLAN 3, and GigabitEthernet 1/0/5 and GigabitEthernet 1/0/6 to VLAN 5.
- The sub-VLANs are isolated at Layer 2 but connected at Layer 3.

Figure 45 Network diagram



Configuration procedure

Create VLAN 10, and configure its VLAN interface IP address as 10.0.0.1/24.

```
<Sysname> system-view
[Sysname] vlan 10
[Sysname-vlan10] quit
[Sysname] interface vlan-interface 10
[Sysname-Vlan-interface10] ip address 10.0.0.1 255.255.255.0
```

Enable local proxy ARP.

```
[Sysname-Vlan-interface10] local-proxy-arp enable
[Sysname-Vlan-interface10] quit
```

```

# Create VLAN 2, and assign GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 to it.
[Sysname] vlan 2
[Sysname-vlan2] port gigabitethernet 1/0/1 gigabitethernet 1/0/2
[Sysname-vlan2] quit

# Create VLAN 3, and assign GigabitEthernet 1/0/3 and GigabitEthernet 1/0/4 to it.
[Sysname] vlan 3
[Sysname-vlan3] port gigabitethernet 1/0/3 gigabitethernet 1/0/4
[Sysname-vlan3] quit

# Create VLAN 5, and assign GigabitEthernet 1/0/5 and GigabitEthernet 1/0/6 to it.
[Sysname] vlan 5
[Sysname-vlan5] port gigabitethernet 1/0/5 gigabitethernet 1/0/6
[Sysname-vlan5] quit

# Configure VLAN 10 as the super VLAN, and configure VLAN 2, VLAN 3, and VLAN 5 as its
sub-VLANs.
[Sysname] vlan 10
[Sysname-vlan10] supervlan
[Sysname-vlan10] subvlan 2 3 5
[Sysname-vlan10] quit
[Sysname] quit

```

Verifying the configurations

```

# Display information about VLAN 10, the super VLAN, to verify the configuration.

```

```

<Sysname> display supervlan
  SuperVLAN ID : 10
  SubVLAN ID : 2-3 5

VLAN ID: 10
VLAN Type: static
It is a Super VLAN.
Route Interface: configured
IPv4 Address: 10.0.0.1
IPv4 Subnet Mask: 255.255.255.0
Description: VLAN 0010
Name: VLAN 0010
Tagged Ports: none
Untagged Ports: none

VLAN ID: 2
VLAN Type: static
It is a Sub VLAN.
Route Interface: configured
IPv4 Address: 10.0.0.1
IPv4 Subnet Mask: 255.255.255.0
Description: VLAN 0002
Name: VLAN 0002
Tagged Ports: none
Untagged Ports:

```

GigabitEthernet1/0/1 GigabitEthernet1/0/2

VLAN ID: 3

VLAN Type: static

It is a Sub VLAN.

Route Interface: configured

IPv4 Address: 10.0.0.1

IPv4 Subnet Mask: 255.255.255.0

Description: VLAN 0003

Name: VLAN 0003

Tagged Ports: none

Untagged Ports:

 GigabitEthernet1/0/3 GigabitEthernet1/0/4

VLAN ID: 5

VLAN Type: static

It is a Sub VLAN.

Route Interface: configured

IPv4 Address: 10.0.0.1

IPv4 Subnet Mask: 255.255.255.0

Description: VLAN 0005

Name: VLAN 0005

Tagged Ports: none

Untagged Ports:

 GigabitEthernet1/0/5 GigabitEthernet1/0/6

Configuring an isolate-user-VLAN

Overview

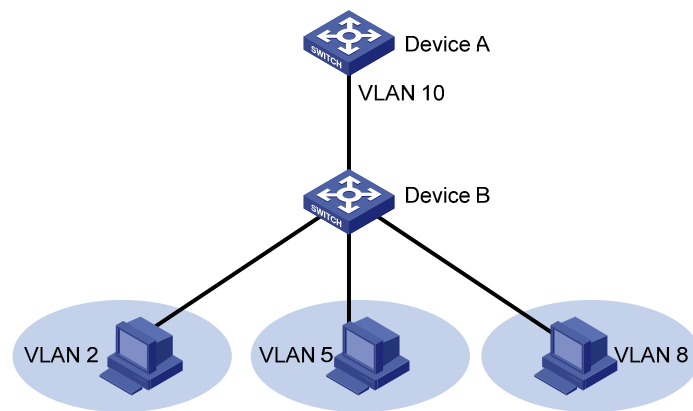
An isolate-user-VLAN uses a two-tier VLAN structure. In this approach, the following types of VLANs, isolate-user-VLAN and secondary VLAN, are configured on the same device.

The following are the characteristics of the isolate-user-VLAN implementation:

- Isolate-user-VLANs are mainly used for upstream data exchange. An isolate-user-VLAN can be associated with multiple secondary VLANs. As the upstream device identifies only the isolate-user-VLAN and not the secondary VLANs, network configuration is simplified and VLAN resources are saved.
- You can isolate the Layer 2 traffic of different users by assigning the ports connected to them to different secondary VLANs. To enable communication between secondary VLANs associated with the same isolate-user-VLAN, you can enable local proxy ARP on the upstream device (for example, Device A in [Figure 46](#)) to realize Layer 3 communication between the secondary VLANs.

As shown in [Figure 46](#), the isolate-user-VLAN function is enabled on Device B. VLAN 10 is the isolate-user-VLAN. VLAN 2, VLAN 5, and VLAN 8 are secondary VLANs associated with VLAN 10 and are invisible to Device A.

Figure 46 An isolate-user-VLAN example



To configure an isolate-user-VLAN, complete the following tasks:

1. Configure the isolate-user-VLAN.
2. Configure the secondary VLANs.
3. Configure uplink and downlink ports:
 - Configure the uplink ports, for example, the port connecting Device B to Device A in [Figure 46](#), to operate in **promiscuous** mode in the specified VLAN, so that the uplink ports can be added to the specified isolate-user-VLAN and the secondary VLANs associated with the isolate-user-VLAN synchronously.
 - Configure the downlink ports, for example, the ports connecting Device B to hosts in [Figure 46](#), to operate in host mode, so that the downlink ports can be added to the isolate-user-VLAN associated with the secondary VLAN synchronously.

- For more information about the promiscuous and host mode commands, see *Layer 2—LAN Switching Command Reference*.
- 4. Associate the isolate-user-VLAN with the specified secondary VLANs.

Configuration restrictions and guidelines

- To enable users in the isolate-user-VLAN to communicate with other networks at Layer 3, follow these steps:
 - a. Configure VLAN interfaces for the isolate-user-VLAN and the secondary VLANs, and configure the gateway IP address for the isolate-user-VLAN interface (you do not need to configure IP addresses for the secondary VLAN interfaces).
 - b. You must configure the **isolated-vlan enable** command for at least one secondary VLAN to isolate the ports in the secondary VLAN.
- The dynamic MAC addresses entries learned in the isolate-user-VLAN are automatically synchronized to all the secondary VLANs, and the dynamic MAC address entries learned in a secondary VLAN are automatically synchronized to the isolate-user-VLAN. Static MAC address entries cannot be automatically synchronized. If you have configured static MAC address entries in the isolate-user-VLAN, you should also configure the same static MAC address entries in the secondary VLANs to avoid broadcasts, and vice versa.
- You cannot configure the member port of a service loopback group as the uplink or downlink port of an isolate-user-VLAN. For more information about the service loopback group, see "[Configuring a service loopback group \(available only on the 5500 EI\)](#)."

Configuration procedure

To configure an isolate-user-VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a VLAN and enter VLAN view.	vlan <i>vlan-id</i>	N/A
3. Configure the VLAN as an isolate-user-VLAN.	isolate-user-vlan enable	Not configured by default.
4. Return to system view.	quit	N/A
5. Create secondary VLANs.	vlan { <i>vlan-id1</i> [<i>to</i> <i>vlan-id2</i>] all }	N/A
6. Configure Layer 2 isolation between ports in the same secondary VLAN.	isolated-vlan enable	Optional. By default, ports in the same secondary VLAN can communicate with one another at Layer 2. This configuration takes effect only after you configure the downlink ports to operate in host mode and associate the secondary VLANs with an isolate-user-VLAN.
7. Return to system view.	quit	N/A

Step	Command	Remarks
8. Configure the uplink port.	<ul style="list-style-type: none"> a. Enter Layer 2 Ethernet or aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> Or interface bridge-aggregation <i>interface-number</i> b. Configure the port to operate in promiscuous mode in a specific VLAN: port isolate-user-vlan <i>vlan-id</i> promiscuous 	N/A
9. Return to system view.	quit	N/A
10. Configure the downlink port.	<ul style="list-style-type: none"> a. Enter Layer 2 Ethernet or aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> Or interface bridge-aggregation <i>interface-number</i> b. Configure the link type of the port: port link-type { access hybrid trunk } c. Assign ports to the secondary VLAN: Access port: port access vlan <i>vlan-id</i> Hybrid port: port hybrid vlan <i>vlan-id-list</i> { tagged untagged } Trunk port: port trunk permit vlan { <i>vlan-id-list</i> all } d. Configure the port to operate in host mode: port isolate-user-vlan host 	N/A
11. Return to system view.	quit	N/A
12. Associate the isolate-user-VLAN with the specified secondary VLANs.	isolate-user-vlan <i>isolate-user-vlan-id</i> secondary <i>secondary-vlan-id</i> [to <i>secondary-vlan-id</i>]	Not configured by default.

Displaying and maintaining isolate-user-VLAN

Task	Command	Remarks
Display the mapping between an isolate-user-VLAN and its secondary VLANs.	display isolate-user-vlan [<i>isolate-user-vlan-id</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

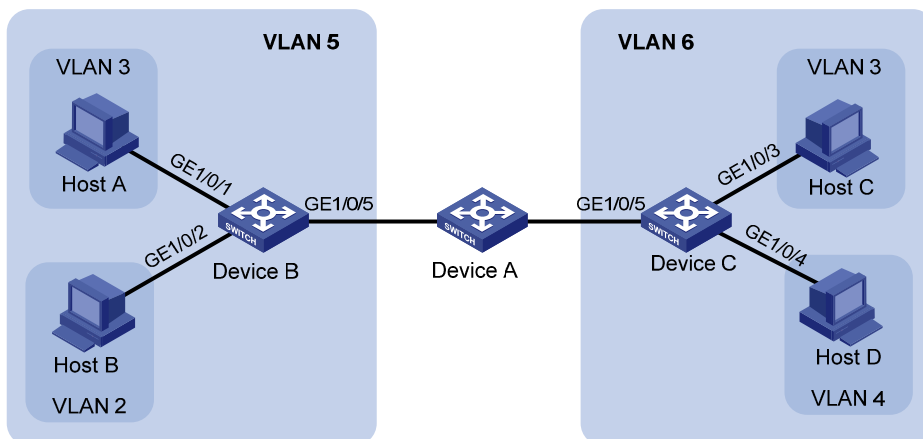
Isolate-user-VLAN configuration example

Network requirements

As shown in [Figure 47](#):

- Connect Device A to downstream devices Device B and Device C.
- Configure VLAN 5 on Device B as an isolate-user-VLAN, assign the uplink port GigabitEthernet 1/0/5 to VLAN 5, and associate VLAN 5 with secondary VLANs VLAN 2 and VLAN 3. Assign GigabitEthernet 1/0/2 to VLAN 2 and GigabitEthernet 1/0/1 to VLAN 3.
- Configure VLAN 6 on Device C as an isolate-user-VLAN, assign the uplink port GigabitEthernet 1/0/5 to VLAN 6, and associate VLAN 6 with secondary VLANs VLAN 3 and VLAN 4. Assign GigabitEthernet 1/0/3 to VLAN 3 and GigabitEthernet 1/0/4 to VLAN 4.
- As far as Device A is concerned, Device B only has VLAN 5 and Device C only has VLAN 6.

Figure 47 Network diagram



Configuration procedure

The following part provides only the configuration on Device B and Device C.

1. Configure Device B:

Configure the isolate-user-VLAN.

```
<DeviceB> system-view
[DeviceB] vlan 5
[DeviceB-vlan5] isolate-user-vlan enable
[DeviceB-vlan5] quit
```

Create secondary VLANs.

```
[DeviceB] vlan 2 to 3
```

Configure the uplink port GigabitEthernet 1/0/5 to operate in promiscuous mode in VLAN 5.

```
[DeviceB] interface gigabitethernet 1/0/5
[DeviceB-GigabitEthernet1/0/5] port isolate-user-vlan 5 promiscuous
```

```
[DeviceB-GigabitEthernet1/0/5] quit
# Assign downlink ports GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 to VLAN 3 and
VLAN 2, respectively, and configure the ports to operate in host mode.
[DeviceB] interface gigabitethernet 1/0/1
[DeviceB-GigabitEthernet1/0/1] port access vlan 3
[DeviceB-GigabitEthernet1/0/1] port isolate-user-vlan host
[DeviceB-GigabitEthernet1/0/1] quit
[DeviceB] interface gigabitethernet 1/0/2
[DeviceB-GigabitEthernet1/0/2] port access vlan 2
[DeviceB-GigabitEthernet1/0/2] port isolate-user-vlan host
[DeviceB-GigabitEthernet1/0/2] quit
# Associate the isolate-user-VLAN with the secondary VLANs.
[DeviceB] isolate-user-vlan 5 secondary 2 to 3
```

2. Configure Device C:

Configure the isolate-user-VLAN.

```
<DeviceC> system-view
[DeviceC] vlan 6
[DeviceC-vlan6] isolate-user-vlan enable
[DeviceC-vlan6] quit
```

Create secondary VLANs.

```
[DeviceC] vlan 3 to 4
```

Configure the uplink port GigabitEthernet 1/0/5 to operate in promiscuous mode in VLAN 6.

```
[DeviceC] interface gigabitethernet 1/0/5
[DeviceC-GigabitEthernet1/0/5] port isolate-user-vlan 6 promiscuous
[DeviceC-GigabitEthernet1/0/5] quit
```

Configure downlink ports GigabitEthernet 1/0/3 and GigabitEthernet 1/0/4 to VLAN 3 and VLAN 4, respectively, and configure the ports to operate in host mode.

```
[DeviceC] interface gigabitethernet 1/0/3
[DeviceC-GigabitEthernet1/0/3] port access vlan 3
[DeviceC-GigabitEthernet1/0/3] port isolate-user-vlan host
[DeviceC-GigabitEthernet1/0/3] quit
[DeviceC] interface gigabitethernet 1/0/4
[DeviceC-GigabitEthernet1/0/4] port access vlan 4
[DeviceC-GigabitEthernet1/0/4] port isolate-user-vlan host
[DeviceC-GigabitEthernet1/0/4] quit
```

Associate the isolate-user-VLAN with the secondary VLANs.

```
[DeviceC] isolate-user-vlan 6 secondary 3 to 4
```

Verifying the configuration

Display the isolate-user-VLAN configuration on Device B.

```
[DeviceB] display isolate-user-vlan
Isolate-user-VLAN VLAN ID : 5
Secondary VLAN ID : 2-3

VLAN ID: 5
VLAN Type: static
Isolate-user-VLAN type : isolate-user-VLAN
```

Route Interface: not configured

Description: VLAN 0005

Name: VLAN 0005

Tagged Ports: none

Untagged Ports:

GigabitEthernet1/0/1

GigabitEthernet1/0/2

GigabitEthernet1/0/5

VLAN ID: 2

VLAN Type: static

Isolate-user-VLAN type : secondary

Route Interface: not configured

Description: VLAN 0002

Name: VLAN 0002

Tagged Ports: none

Untagged Ports:

GigabitEthernet1/0/2

GigabitEthernet1/0/5

VLAN ID: 3

VLAN Type: static

Isolate-user-VLAN type : secondary

Route Interface: not configured

Description: VLAN 0003

Name: VLAN 0003

Tagged Ports: none

Untagged Ports:

GigabitEthernet1/0/1

GigabitEthernet1/0/5

Configuring a voice VLAN

Overview

A voice VLAN is configured for voice traffic. After assigning the ports that connect to voice devices to a voice VLAN, the system automatically configures quality of service (QoS) parameters for voice traffic, to improve the transmission priority of voice traffic and ensure voice quality.

Common voice devices include IP phones and integrated access devices (IADs). Only IP phones are used in the voice VLAN configuration examples in this document.

OUI addresses

A device determines whether a received packet is a voice packet by evaluating its source MAC address. A packet whose source MAC address complies with the Organizationally Unique Identifier (OUI) address of the voice device is regarded as voice traffic.

You can remove the default OUI address of a device manually and then add new ones manually. You can configure the OUI addresses of a device in advance or use the default OUI addresses. [Table 15](#) lists the default OUI address for each vendor's devices.

Table 15 The default OUI addresses of different vendors

Number	OUI address	Vendor
1	0001-E300-0000	Siemens phone
2	0003-6B00-0000	Cisco phone
3	0004-0D00-0000	Avaya phone
4	00D0-1E00-0000	Pingtel phone
5	0060-B900-0000	Philips/NEC phone
6	00E0-7500-0000	Polycom phone
7	00E0-BB00-0000	3Com phone

In general, as the first 24 bits of a MAC address (in binary format), an OUI address is a globally unique identifier that IEEE assigns to a vendor. In this document, however, OUI addresses are addresses that the system uses to determine whether a received packet is a voice packet. They are the results of the AND operation of the arguments *mac-address* and *oui-mask* in the **voice vlan mac-address** command.

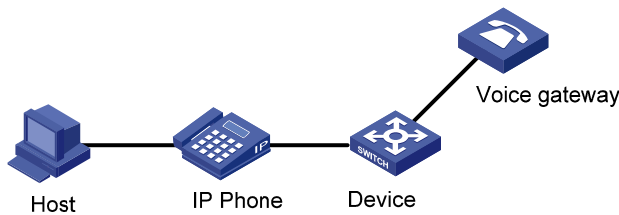
Voice VLAN assignment modes

A port can be assigned to a voice VLAN in one of the following modes:

- **Automatic mode**—The system matches the source MAC address carried in the untagged packets sent when an IP phone is powered on against the device's OUI addresses. If the system finds a match, it automatically assigns the receiving port to the voice VLAN, issues ACL rules, and configures the packet precedence. You can configure a voice VLAN aging time on the device. The system will remove a port from the voice VLAN if no packet is received from the port during the aging time. The system automatically assigns ports to, or removes ports from, a voice VLAN.

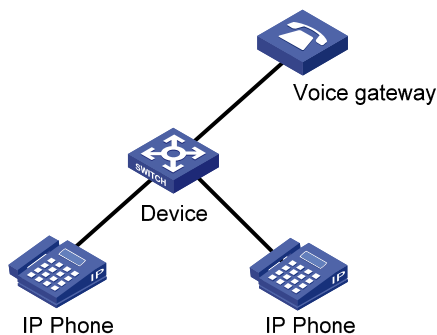
Automatic mode is suitable for scenarios where PCs and IP phones connected in series access the network through the device and ports on the device transmit both voice traffic and data traffic at the same time, as shown in Figure 48. When the voice VLAN works normally, when the system reboots, the system reassigns ports in automatic voice VLAN assignment mode to the voice VLAN after the reboot, ensuring that existing voice connections can work normally. In this case, voice traffic streams do not trigger port assignment to the voice VLAN.

Figure 48 PCs and IP phones connected in series access the network



- Manual mode**—You must manually assign an IP phone accessing port to a voice VLAN. Then, the system matches the source MAC addresses carried in the packets against the device’s OUI addresses. If the system finds a match, it issues ACL rules and configures the packet precedence. In this mode, you must manually assign ports to, or remove ports from, a voice VLAN. Manual mode is suitable for scenarios where only IP phones access the network through the device and ports on the device transmit only voice traffic, as shown in Figure 49. In this mode, ports assigned to a voice VLAN transmit voice traffic exclusively, which prevents the impact of data traffic on the transmission of voice traffic.

Figure 49 Only IP phones access the network



Both modes forward tagged packets according to their tags.

Table 16 and Table 17 list the configurations required for ports of different link types to support tagged or untagged voice traffic sent from IP phones when different voice VLAN assignment modes are configured.

- IP phones send tagged voice traffic

Table 16 Required configurations on ports of different link types for them to support tagged voice traffic

Port link type	Voice VLAN assignment mode	Support for tagged voice traffic	Configuration requirements
Access	Automatic	No	N/A
	Manual		
Trunk	Automatic	Yes	The PVID of the port cannot be the voice VLAN.

Port link type	Voice VLAN assignment mode	Support for tagged voice traffic	Configuration requirements
	Manual		The PVID of the port cannot be the voice VLAN. Configure the port to permit packets of the voice VLAN to pass through.
	Automatic		The PVID of the port cannot be the voice VLAN.
Hybrid	Manual	Yes	The PVID of the port cannot be the voice VLAN. Configure the port to permit packets of the voice VLAN to pass through tagged.

- IP phones send untagged voice traffic

When IP phones send untagged voice traffic, you can only configure the voice traffic receiving ports on the device to operate in manual voice VLAN assignment mode.

Table 17 Required configurations on ports of different link types for them to support tagged voice traffic

Port link type	Voice VLAN assignment mode	Support for untagged voice traffic	Configuration requirements
Access	Automatic	No	N/A
	Manual	Yes	Configure the PVID of the port as the voice VLAN.
Trunk	Automatic	No	N/A
	Manual	Yes	Configure the PVID of the port as the voice VLAN and assign the port to the voice VLAN.
Hybrid	Automatic	No	N/A
	Manual	Yes	Configure the PVID of the port as the voice VLAN and configure the port to permit packets of the voice VLAN to pass through untagged.

When you configure the voice VLAN assignment modes, follow these guidelines:

- If an IP phone sends tagged voice traffic and its accessing port is configured with 802.1X authentication and guest VLAN, assign different VLAN IDs for the voice VLAN, the PVID of the connecting port, and the 802.1X guest VLAN.
- If an IP phone sends untagged voice traffic, to implement the voice VLAN feature, you must configure the PVID of the IP phone's accessing port as the voice VLAN. As a result, you cannot implement 802.1X authentication.
- The PVID is VLAN 1 for all ports by default. You can configure the PVID of a port and assign a port to certain VLANs by using commands. For more information, see "[Configuring VLANs.](#)"
- Use the **display interface** command to display the PVID of a port and the VLANs to which the port is assigned.

Security mode and normal mode of voice VLANs

Depending on their inbound packet filtering mechanisms, voice VLAN-enabled ports operate in the following modes:

- **Normal mode**—Voice VLAN-enabled ports receive packets that carry the voice VLAN tag, and forward packets in the voice VLAN without comparing their source MAC addresses against the OUI addresses configured for the device. If the PVID of the port is the voice VLAN and the port operates

in manual VLAN assignment mode, the port forwards all received untagged packets in the voice VLAN. In normal mode, voice VLANs are vulnerable to traffic attacks. Malicious users might send large quantities of forged voice packets to consume the voice VLAN bandwidth, affecting normal voice communication.

- **Security mode**—Only voice packets whose source MAC addresses match the recognizable OUI addresses can pass through the voice VLAN-enabled inbound port, but all other packets are dropped.

In a safe network, you can configure the voice VLANs to operate in normal mode, reducing the consumption of system resources due to source MAC addresses checking.



TIP:

HP does not recommend you transmit both voice traffic and non-voice traffic in a voice VLAN. If you must transmit both voice traffic and nonvoice traffic, make sure that the voice VLAN security mode is disabled.

Table 18 How a voice VLAN-enabled port processes packets in security and normal mode

Voice VLAN mode	Packet type	Packet processing mode
Security mode	Untagged packets	If the source MAC address of a packet matches an OUI address configured for the device, it is forwarded in the voice VLAN; otherwise, it is dropped.
	Packets that carry the voice VLAN tag	
	Packets that carry other tags	Forwarded or dropped depending on whether the port allows packets of these VLANs to pass through
Normal mode	Untagged packets	The port does not determine the source MAC addresses of inbound packets. In this way, both voice traffic and non-voice traffic can be transmitted in the voice VLAN.
	Packets that carry the voice VLAN tag	
	Packets that carry other tags	Forwarded or dropped depending on whether the port allows packets of these VLANs to pass through

Configuration prerequisites

Before you configure a voice VLAN, complete the following tasks:

- Create a VLAN.
- Configure QoS priority settings for voice VLAN traffic on an interface before you enable voice VLAN on the interface.

If the configuration order is reversed, your priority configuration will fail. For more information, see "[Configuring QoS priority settings for voice traffic on an interface.](#)"

- Configure the voice VLAN assignment mode.

For more information, see "[Configuring a port to operate in automatic voice VLAN assignment mode](#)" and "[Configuring a port to operate in manual voice VLAN assignment mode.](#)"

Configuring QoS priority settings for voice traffic on an interface

In voice VLAN applications, you can improve the quality of voice traffic by configuring the appropriate QoS priority settings, including the Class of Service (CoS) and Differentiated Services Code Point (DSCP) values, for voice traffic. Voice traffic carries its own QoS priority settings. You can configure the device either to modify or not to modify the QoS priority settings carried by incoming voice traffic.

Configuration restrictions and guidelines

Configure the QoS priority settings for voice traffic on an interface before you enable voice VLAN on the interface. If the configuration order is reversed, your priority trust setting will fail.

Configuration procedure

To configure QoS priority settings for voice traffic:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the interface to trust the QoS priority settings in incoming voice traffic, but not to modify the CoS and DSCP values marked for incoming traffic of the voice VLAN.	voice vlan qos trust	Use either command. By default, an interface modifies the CoS value and the DSCP value marked for voice VLAN traffic into 6 and 46, respectively.
4. Configure the interface to modify the CoS and DSCP values marked for incoming traffic of the voice VLAN into specified values.	voice vlan qos <i>cos-value</i> <i>dscp-value</i>	The voice vlan qos command and the voice vlan qos trust command can overwrite each other, whichever is configured last.

Configuring a port to operate in automatic voice VLAN assignment mode

To set a port to operate in automatic voice VLAN assignment mode:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Set the voice VLAN aging time.	voice vlan aging <i>minutes</i>	Optional. By default, the aging time of a voice VLAN is 1440 minutes. The voice VLAN aging time configuration is only applicable on ports in automatic voice VLAN assignment mode.
3. Enable the voice VLAN security mode.	voice vlan security enable	Optional. By default, the voice VLAN security mode is enabled.
4. Add a recognizable OUI address.	voice vlan mac-address <i>oui mask oui-mask [description text]</i>	Optional. By default, each voice VLAN has default OUI addresses configured. For the default OUI addresses of different vendors, see Table 15 .
5. Enter Ethernet interface view.	interface <i>interface-type interface-number</i>	N/A
6. Configure the port to operate in automatic voice VLAN assignment mode.	voice vlan mode auto	Optional. By default, the automatic voice VLAN assignment mode is enabled. The voice VLAN assignment modes on different ports are independent of one another.
7. Enable the voice VLAN feature.	voice vlan <i>vlan-id</i> enable	By default, the voice VLAN feature is disabled.

NOTE:

A protocol-based VLAN on a hybrid port can process only untagged inbound packets, whereas the voice VLAN in automatic mode on a hybrid port can process only tagged voice traffic. Do not configure a VLAN as both a protocol-based VLAN and a voice VLAN. For more information, see "[Configuring VLANs](#)."

Configuring a port to operate in manual voice VLAN assignment mode

Configuration restrictions and guidelines

- You can configure different voice VLANs on different ports at the same time. However, you can configure one port with only one voice VLAN, and this voice VLAN must be a static VLAN that already exists on the device.
- You cannot enable voice VLAN on the member ports of a link aggregation group. For more information about the member ports, see "[Configuring Ethernet link aggregation](#)."
- To make voice VLAN take effect on a port that is enabled with voice VLAN and operates in manual voice VLAN assignment mode, you must assign the port to the voice VLAN manually.

Configuration procedure

To set a port to operate in manual voice VLAN assignment mode:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable the voice VLAN security mode.	voice vlan security enable	Optional. By default, the voice VLAN security mode is enabled.
3. Add a recognizable OUI address.	voice vlan mac-address <i>oui mask</i> <i>oui-mask</i> [description <i>text</i>]	Optional. By default, each voice VLAN has default OUI addresses configured. For the default OUI addresses of different vendors, see Table 15 .
4. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
5. Configure the port to operate in manual voice VLAN assignment mode.	undo voice vlan mode auto	By default, the manual voice VLAN assignment mode is disabled.
6. Assign the access, trunk, or hybrid port in manual voice VLAN assignment mode to the voice VLAN.	For the configuration procedure, see " Configuring VLANs ."	After you assign an access port to the voice VLAN, the voice VLAN becomes the PVID of the port automatically.
7. Configure the voice VLAN as the PVID of the trunk or hybrid port.	For the configuration procedure, see " Configuring VLANs ."	Optional. This operation is required for untagged inbound voice traffic and prohibited for tagged inbound voice traffic.
8. Enable voice VLAN on the port.	voice vlan <i>vlan-id</i> enable	Disabled by default.

Displaying and maintaining voice VLAN

Task	Command	Remarks
Display the voice VLAN state.	display voice vlan state [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the OUI addresses that the system supports.	display voice vlan oui [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Voice VLAN configuration examples

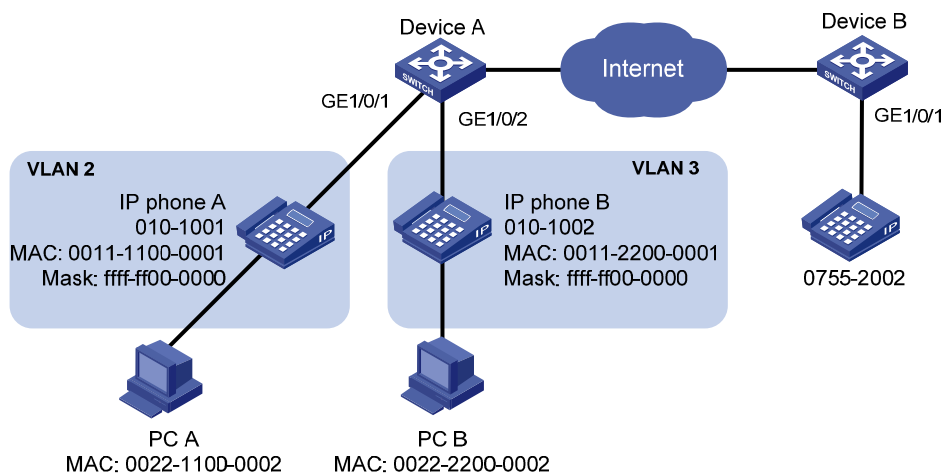
Automatic voice VLAN mode configuration example

Network requirements

As shown in [Figure 50](#),

- The MAC address of IP phone A is 0011-1100-0001. The phone connects to a downstream device named PC A whose MAC address is 0022-1100-0002 and to GigabitEthernet 1/0/1 on an upstream device named Device A.
- The MAC address of IP phone B is 0011-2200-0001. The phone connects to a downstream device named PC B whose MAC address is 0022-2200-0002 and to GigabitEthernet 1/0/2 on Device A.
- Device A uses voice VLAN 2 to transmit voice packets for IP phone A and uses voice VLAN 3 to transmit voice packets for IP phone B.
- Configure GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 to operate in automatic voice VLAN assignment mode. In addition, if one of them has not received any voice packet in 30 minutes, the port is removed from the corresponding voice VLAN automatically.

Figure 50 Network diagram



Configuration procedure

Create VLAN 2 and VLAN 3.

```
<DeviceA> system-view
[DeviceA] vlan 2 to 3
Please wait... Done.
```

Set the voice VLAN aging time to 30 minutes.

```
[DeviceA] voice vlan aging 30
```

Since GigabitEthernet 1/0/1 might receive both voice traffic and data traffic at the same time, to ensure the quality of voice packets and effective bandwidth use, configure voice VLANs to operate in security mode. Configure the voice VLANs to transmit only voice packets. By default, voice VLANs operate in security mode. (Optional)

```
[DeviceA] voice vlan security enable
```

Configure the allowed OUI addresses as MAC addresses prefixed by 0011-1100-0000 or 0011-2200-0000. In this way, Device A identifies packets whose MAC addresses match any of the configured OUI addresses as voice packets.

```
[DeviceA] voice vlan mac-address 0011-1100-0001 mask ffff-ff00-0000 description IP phone A
[DeviceA] voice vlan mac-address 0011-2200-0001 mask ffff-ff00-0000 description IP phone B
```

Configure GigabitEthernet 1/0/1 as a hybrid port.

```
[DeviceA] interface gigabitethernet 1/0/1
```

```

[DeviceA-GigabitEthernet1/0/1] port link-type hybrid

# Configure GigabitEthernet 1/0/1 to operate in automatic voice VLAN assignment mode. By default,
a port operates in automatic voice VLAN assignment mode. (Optional)
[DeviceA-GigabitEthernet1/0/1] voice vlan mode auto

# Configure VLAN 2 as the voice VLAN for GigabitEthernet 1/0/1.
[DeviceA-GigabitEthernet1/0/1] voice vlan 2 enable
[DeviceA-GigabitEthernet1/0/1] quit

# Configure GigabitEthernet 1/0/2.
[DeviceA] interface gigabitethernet 1/0/2
[DeviceA-GigabitEthernet1/0/2] port link-type hybrid
[DeviceA-GigabitEthernet1/0/2] voice vlan mode auto
[DeviceA-GigabitEthernet1/0/2] voice vlan 3 enable

```

Verifying the configurations

Display the OUI addresses, OUI address masks, and description strings.

```

<DeviceA> display voice vlan oui
Oui Address      Mask             Description
0001-e300-0000   ffff-ff00-0000   Siemens phone
0003-6b00-0000   ffff-ff00-0000   Cisco phone
0004-0d00-0000   ffff-ff00-0000   Avaya phone
0011-1100-0000   ffff-ff00-0000   IP phone A
0011-2200-0000   ffff-ff00-0000   IP phone B
0060-b900-0000   ffff-ff00-0000   Philips/NEC phone
00d0-1e00-0000   ffff-ff00-0000   Pingtel phone
00e0-7500-0000   ffff-ff00-0000   Polycom phone
00e0-bb00-0000   ffff-ff00-0000   3com phone

```

Display the states of voice VLANs.

```

<DeviceA> display voice vlan state
Maximum of Voice VLANs: 8
Current Voice VLANs: 2
Voice VLAN security mode: Security
Voice VLAN aging time: 30 minutes
Voice VLAN enabled port and its mode:

```

PORT	VLAN	MODE	COS	DSCP
GigabitEthernet1/0/1	2	AUTO	6	46
GigabitEthernet1/0/2	3	AUTO	6	46

Manual voice VLAN assignment mode configuration example

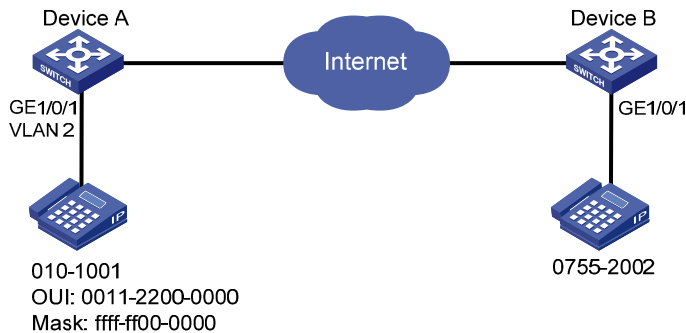
Network requirements

As shown in [Figure 51](#),

- Create VLAN 2 and configure it as a voice VLAN that permits only voice traffic to pass through.
- The IP phones send untagged voice traffic. Configure GigabitEthernet 1/0/1 as a hybrid port.

- Configure GigabitEthernet 1/0/1 to operate in manual voice VLAN assignment mode. Configure GigabitEthernet 1/0/1 to allow voice traffic with an OUI address of 0011-2200-0000, a mask of ffff-ff00-0000, and a description string of **test** to be forwarded in the voice VLAN.

Figure 51 Network diagram



Configuration procedure

Configure the voice VLAN to operate in security mode. A voice VLAN operates in security mode by default. (Optional)

```
<DeviceA> system-view
[DeviceA] voice vlan security enable
```

Add a recognizable OUI address 0011-2200-0000.

```
[DeviceA] voice vlan mac-address 0011-2200-0000 mask ffff-ff00-0000 description test
```

Create VLAN 2.

```
[DeviceA] vlan 2
[DeviceA-vlan2] quit
```

Configure GigabitEthernet 1/0/1 to operate in manual voice VLAN assignment mode.

```
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] undo voice vlan mode auto
```

Configure GigabitEthernet 1/0/1 as a hybrid port.

```
[DeviceA-GigabitEthernet1/0/1] port link-type hybrid
```

Configure the voice VLAN (VLAN 2) as the PVID of GigabitEthernet 1/0/1 and configure GigabitEthernet 1/0/1 to permit the voice traffic of VLAN 2 to pass through untagged.

```
[DeviceA-GigabitEthernet1/0/1] port hybrid pvid vlan 2
[DeviceA-GigabitEthernet1/0/1] port hybrid vlan 2 untagged
```

Enable voice VLAN on GigabitEthernet 1/0/1.

```
[DeviceA-GigabitEthernet1/0/1] voice vlan 2 enable
```

Verifying the configurations

Display the OUI addresses, OUI address masks, and description strings.

```
<DeviceA> display voice vlan oui
```

Oui Address	Mask	Description
0001-e300-0000	ffff-ff00-0000	Siemens phone
0003-6b00-0000	ffff-ff00-0000	Cisco phone
0004-0d00-0000	ffff-ff00-0000	Avaya phone
0011-2200-0000	ffff-ff00-0000	test
00d0-1e00-0000	ffff-ff00-0000	Pingtel phone

```
0060-b900-0000 ffff-ff00-0000 Philips/NEC phone
00e0-7500-0000 ffff-ff00-0000 Polycom phone
00e0-bb00-0000 ffff-ff00-0000 3com phone
```

Display the states of voice VLANs.

```
<DeviceA> display voice vlan state
```

```
Maximum of Voice VLANs: 8
```

```
Current Voice VLANs: 1
```

```
Voice VLAN security mode: Security
```

```
Voice VLAN aging time: 1440 minutes
```

```
Voice VLAN enabled port and its mode:
```

PORT	VLAN	MODE	COS	DSCP

GigabitEthernet1/0/1	2	MANUAL	6	46

Configuring GVRP

The Generic Attribute Registration Protocol (GARP) provides a generic framework for devices in a switched LAN, such as end stations and switches, to register and deregister attribute values. The GARP VLAN Registration Protocol (GVRP) is a GARP application that registers and deregisters VLAN attributes. GVRP uses the operating mechanism of GARP to maintain and propagate dynamic VLAN registration information for GVRP devices on the network.

Overview

GARP

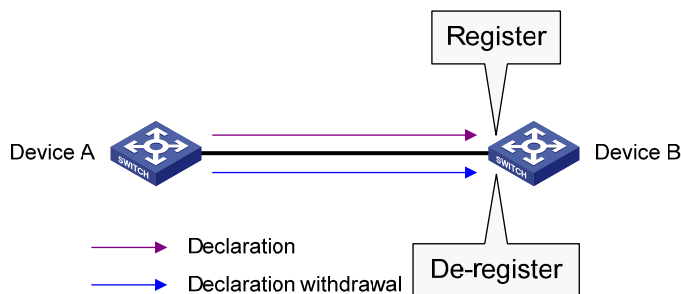
GARP provides a mechanism that allows participants in a GARP application to distribute, propagate, and register—with other participants in a LAN—the attributes specific to the GARP application, such as VLAN or multicast address attributes.

How GARP works

Each port that participates in a GARP application (GVRP, for example) is a GARP participant.

GARP enables GARP participants to propagate attribute values throughout the switched LAN. As shown in Figure 52, a GARP participant registers and deregisters its attribute values with other GARP participants by sending and withdrawing declarations, and registers and deregisters the attribute values of other participants according to the declarations and withdrawals that it has received.

Figure 52 How GARP works



For example, a GVRP-enabled port registers and deregisters VLAN in the following cases.

- When the port receives a VLAN attribute declaration, it registers the VLAN attribute and joins the VLAN.
- When the port receives a VLAN withdrawal, it deregisters the VLAN and leaves the VLAN.

GARP messages

A GARP participant exchanges information with other GARP participants by sending GARP messages: Join, Leave, and LeaveAll. As a GARP application, GVRP also uses GARP messages for information exchange.

- Join messages

A GARP participant sends Join messages when it wishes to declare its attribute values or receives Join messages from other GARP participants.

Join messages fall into JoinEmpty and JoinIn. A GARP participant sends JoinEmpty messages to declare attribute values that it has not registered. It sends JoinIn messages to declare attribute values that it has registered.

- Leave messages

A GARP participant sends Leave messages when it wishes to withdraw declarations of its attribute values (because, for example, it has deregistered its attribute values), or receives Leave messages from other participants.

Leave messages fall into LeaveEmpty and LeaveIn. A GARP participant sends LeaveEmpty messages to withdraw declarations of the attribute values that it has not registered. It sends LeaveIn messages to withdraw declarations of the attribute values that it has registered.

- LeaveAll messages

A GARP participant sends a LeaveAll message when it declares that it is deregistering all attribute values or receives LeaveAll messages from other participants. If any participants want to maintain the registration for a particular attribute value, they must send a Join message.

GARP timers

HP's implementation of GARP uses the following timers to control GARP message transmission:

- Hold timer

The Hold timer sets the delay that a GARP participant waits before sending a Join or Leave message.

When an attribute value changes or a Join or Leave message arrives, the GARP participant does not send the message immediately. Rather, it assembles Join and Leave messages in the least number of GARP PDUs, and sends them out when the Hold timer expires. This timer reduces the number of GARP PDUs and saves bandwidth.

- Join timer

A GARP participant might declare an attribute twice to ensure reliable transmission. The Join timer sets the interval between the two declarations.

A GARP participant starts a Join timer when it declares an attribute value or receives a JoinIn message for the attribute value. If the GARP participant does not receive any declaration for the attribute value when the Join timer expires, it re-declares the attribute value.

Because all attributes of a GARP participant share the same Join timer, you must set the Join timer long enough so that all attributes can be sent out in one declaration.

- Leave timer

A GARP participant starts a Leave timer when it receives a Leave message for an attribute value. If the GARP participant receives no Join message for the attribute value before the timer expires, it deregisters the attribute value.

- LeaveAll timer

When a GARP application is enabled, a LeaveAll timer starts. The GARP participant sends a LeaveAll message when the timer expires. Then, the LeaveAll timer restarts to begin a new cycle. The LeaveAll timer and all other GARP timers also restart when the GARP participant receives a LeaveAll message.

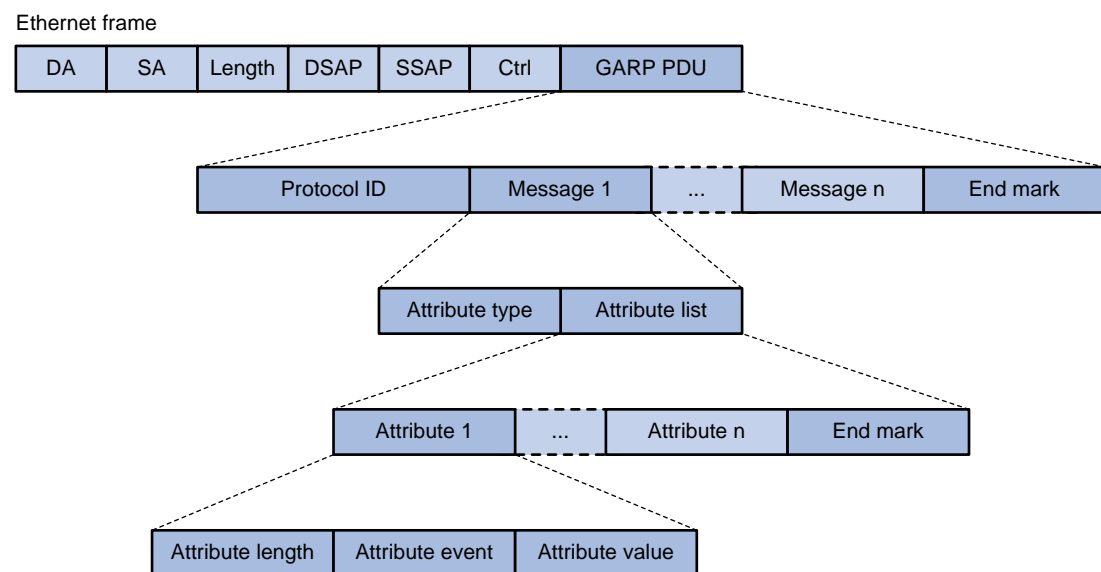
When you configure GARP timers, follow these guidelines:

- The settings of GARP timers apply to all GARP applications, such as GVRP, on a LAN.

- On a GARP-enabled network, each port maintains its own Hold, Join, and Leave timers, but only one LeaveAll timer is maintained on each device. This LeaveAll timer applies to all ports on the device.
- The value ranges for the Hold, Join, Leave, and LeaveAll timers are dependent on one another. See [Table 20](#) for their dependencies.
- Set the LeaveAll timer greater than any Leave timer and not smaller than its default value, 1000 centiseconds. Each time a LeaveAll timer expires, a network-wide re-join occurs.
- A device can send LeaveAll messages at the interval set by its LeaveAll timer or the LeaveAll timer of another device on the network, whichever is smaller. This is because each time a device on the network receives a LeaveAll message, it resets its LeaveAll timer.

GARP PDU format

Figure 53 GARP PDU format



As shown in [Figure 53](#), GARP PDUs are encapsulated in IEEE 802.3 Ethernet frames.

Table 19 GARP PDU fields

Field	Description	Value
Protocol ID	Protocol identifier for GARP	0x0001
Message	One or multiple messages, each of which contains an attribute type and an attribute list	N/A
End mark	Indicates the end of a GARP PDU	0x00
Attribute type	Defined by the GARP application	0x01 for GVRP, which indicates the VLAN ID attribute
Attribute list	Contains one or multiple attributes	N/A
Attribute	Consists of an attribute length, an attribute event, and an attribute value	N/A
Attribute length	Length of an attribute, inclusive of the attribute length field	2 to 255 (in bytes)

Field	Description	Value
Attribute event	Event that the attribute describes	<ul style="list-style-type: none"> • 0x00—LeaveAll event • 0x01—JoinEmpty event • 0x02—JoinIn event • 0x03—LeaveEmpty event • 0x04—LeaveIn event • 0x05—Empty event
Attribute value	Attribute value	VLAN ID for GVRP If the value of the attribute event field is 0x00 (LeaveAll event), the attribute value field is invalid.

The destination MAC addresses of GARP messages are multicast MAC addresses, and vary with GARP applications. For example, the destination MAC address of GVRP is 01-80-C2-00-00-21.

GVRP

GVRP overview

As a GARP application, GVRP uses the operating mechanism of GARP to maintain and propagate dynamic VLAN registrations throughout a switched LAN.

In a switched LAN, each GVRP-enabled switch sends and receives VLAN declarations and withdrawals from other GVRP-enabled switches, and dynamically updates its local database, including active VLAN members and through which port each VLAN member can be reached. This makes sure that all GVRP-enabled switches in a LAN maintain the same VLAN information.

The VLAN information propagated by GVRP includes not only manually configured static VLAN information but also dynamic VLAN information from other switches.

GVRP registration modes

GVRP is available on trunk ports. It provides the following registration modes:

- **Normal mode**—Performs dynamic VLAN registrations and deregistrations on the trunk port, and sends declarations and withdrawals for dynamic and static VLANs. VLANs manually configured are called static VLANs, and VLANs created by GVRP are called dynamic VLANs.
- **Fixed mode**—Disables the trunk port to register or withdraw dynamic VLAN information, but allows the port to send declarations for static VLANs. A trunk port in this mode carries only static VLANs, even if it has been assigned to all VLANs.
- **Forbidden mode**—Disables the trunk port to register or withdraw dynamic VLAN information, and allows the port to send declarations only for VLAN 1. A trunk port in this mode carries only VLAN 1 even if it has been assigned to any other VLANs.

Protocols and standards

IEEE 802.1Q, *Virtual Bridged Local Area Networks*

GVRP configuration task list

When you configure GVRP, follow these guidelines:

- GVRP configuration made in Ethernet interface view or Layer 2 aggregate interface view takes effect on the current interface only; GVRP configuration made in port group view takes effect on all the member ports in the group.
- GVRP configuration made on a member port in an aggregation group takes effect only after the port is removed from the aggregation group.

Complete these tasks to configure GVRP:

Task	Remarks
Configuring GVRP functions	Required
Configuring the GARP timers	Optional

Configuring GVRP functions

Before enabling GVRP on a port, you must enable GVRP globally. In addition, you can configure GVRP only on trunk ports, and you must assign the involved trunk ports to all dynamic VLANs.

Configuration restrictions and guidelines

- GVRP is mutually exclusive with service loopback.
- GVRP can work with STP, RSTP, or MSTP CIST but not PVST. When GVRP runs on the CIST, blocked ports on the CIST cannot receive or send GVRP packets. For more information about STP, RSTP, MSTP CIST, and PVST, see "[Configuring spanning tree protocols.](#)"
- Do not enable both GVRP and remote port mirroring. Otherwise, GVRP may register the remote probe VLAN to unexpected ports, resulting in undesired duplicates to be received by the monitor port. For more information about port mirroring, see *Network Management and Monitoring Configuration Guide*.
- Enabling GVRP on a Layer 2 aggregate interface enables both the aggregate interface and all selected member ports in the corresponding link aggregation group to participate in dynamic VLAN registration and deregistration.

Configuration procedure

To configure GVRP functions on a trunk port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable GVRP globally.	gvrp	Globally disabled by default.
3. Enter Ethernet interface view, Layer 2 aggregate interface view, or port group view.	<ul style="list-style-type: none"> • Enter Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> • Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.

Step	Command	Remarks
4. Configure the link type of the ports as trunk.	port link-type trunk	Access by default. For more information about the port link-type trunk command, see <i>Layer 2—LAN Switching Command Reference</i> .
5. Assign the trunk ports to all VLANs.	port trunk permit vlan all	By default, a trunk port is assigned to VLAN 1 only. For more information about the port trunk permit vlan all command, see <i>Layer 2—LAN Switching Command Reference</i> .
6. Enable GVRP on the ports.	gvrp	Disabled by default.
7. Configure the GVRP registration mode on the port.	gvrp registration { fixed forbidden normal }	Optional. normal by default.

Configuring the GARP timers

To configure the GARP timers:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the GARP LeaveAll timer.	garp timer leaveall timer-value	Optional. 1000 centiseconds by default. The LeaveAll timer applies to all ports.
3. Enter Ethernet interface view, Layer 2 aggregate interface view, or port group view.	<ul style="list-style-type: none"> Enter Ethernet interface view or Layer 2 aggregate interface view: interface interface-type interface-number Enter port group view: port-group manual port-group-name 	Use either command.
4. Configure the Hold timer.	garp timer hold timer-value	Optional. 10 centiseconds by default.
5. Configure the Join timer.	garp timer join timer-value	Optional. 20 centiseconds by default.
6. Configure the Leave timer.	garp timer leave timer-value	Optional. 60 centiseconds by default.

As shown in [Table 20](#), the value ranges for GARP timers are dependent on one another; use the following guidelines to configure GARP timers:

- If you want to set a value beyond the value range for a timer, you can change the value range by tuning the value of another related timer.

- If you want to restore the default settings of the timers, restore the Hold timer first, followed by the Join, Leave, and LeaveAll timers.

Table 20 Dependencies of the GARP timers

Timer	Lower limit	Upper limit
Hold	10 centiseconds	No greater than half of the Join timer
Join	No less than twice the Hold timer	Less than half of the Leave timer
Leave	Greater than twice the Join timer	Less than the LeaveAll timer
LeaveAll	Greater than the Leave timer	32,765 centiseconds

NOTE:

To keep the dynamic VLANs learned through GVRP stable, do not set the LeaveAll timer smaller than its default value, 1000 centiseconds.

Displaying and maintaining GVRP

Task	Command	Remarks
Display statistics about GARP on ports.	display garp statistics [interface <i>interface-list</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display GARP timers on ports.	display garp timer [interface <i>interface-list</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the local VLAN information that GVRP maintains on ports.	display gvrp local-vlan interface <i>interface-type interface-number</i> [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the current GVRP state in the specified VLANs on ports.	display gvrp state interface <i>interface-type interface-number</i> vlan <i>vlan-id</i> [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display GVRP statistics on ports.	display gvrp statistics [interface <i>interface-list</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the global GVRP state.	display gvrp status [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the information about dynamic VLAN operations on ports.	display gvrp vlan-operation interface <i>interface-type interface-number</i> [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear the GARP statistics on ports.	reset garp statistics [interface <i>interface-list</i>]	Available in user view

GVRP configuration examples

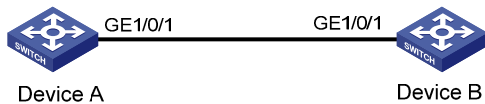
GVRP normal registration mode configuration example

Network requirements

As shown in [Figure 54](#):

- Device A and Device B are connected through their ports GigabitEthernet 1/0/1.
- Enable GVRP and configure the normal registration mode on ports to enable the registration and deregistration of dynamic and static VLAN information between the two devices.

Figure 54 Network diagram



Configuration procedure

1. Configure Device A:

Enable GVRP globally.

```
<DeviceA> system-view
```

```
[DeviceA] gvrp
```

Configure port GigabitEthernet 1/0/1 as a trunk port, and assign it to all VLANs.

```
[DeviceA] interface gigabitethernet 1/0/1
```

```
[DeviceA-GigabitEthernet1/0/1] port link-type trunk
```

```
[DeviceA-GigabitEthernet1/0/1] port trunk permit vlan all
```

Enable GVRP on trunk port GigabitEthernet 1/0/1.

```
[DeviceA-GigabitEthernet1/0/1] gvrp
```

```
[DeviceA-GigabitEthernet1/0/1] quit
```

Create VLAN 2 (a static VLAN).

```
[DeviceA] vlan 2
```

```
[DeviceA-vlan2] quit
```

2. Configure Device B:

Enable GVRP globally.

```
<DeviceB> system-view
```

```
[DeviceB] gvrp
```

Configure port GigabitEthernet 1/0/1 as a trunk port, and assign it to all VLANs.

```
[DeviceB] interface gigabitethernet 1/0/1
```

```
[DeviceB-GigabitEthernet1/0/1] port link-type trunk
```

```
[DeviceB-GigabitEthernet1/0/1] port trunk permit vlan all
```

Enable GVRP on trunk port GigabitEthernet 1/0/1.

```
[DeviceB-GigabitEthernet1/0/1] gvrp
```

```
[DeviceB-GigabitEthernet1/0/1] quit
```

Create VLAN 3 (a static VLAN).

```
[DeviceB] vlan 3
```

```
[DeviceB-vlan3] quit
```

3. Verify the configuration:

Use the **display gvrp local-vlan** command to display the local VLAN information that GVRP maintains on ports. For example:

Display the local VLAN information that GVRP maintains on port GigabitEthernet 1/0/1 of Device A.

```
[DeviceA] display gvrp local-vlan interface gigabitethernet 1/0/1
```

```
Following VLANs exist in GVRP local database:
```

```
1(default),2-3
```

According to the output, information about VLAN 1, static VLAN information of VLAN 2 on the local device, and dynamic VLAN information of VLAN 3 on Device B are all registered through GVRP.

Display the local VLAN information that GVRP maintains on port GigabitEthernet 1/0/1 of Device B.

```
[DeviceB] display gvrp local-vlan interface gigabitethernet 1/0/1
```

Following VLANs exist in GVRP local database:

```
1(default),2-3
```

According to the output, information about VLAN 1, static VLAN information of VLAN 3 on the local device, and dynamic VLAN information of VLAN 2 on Device A are all registered through GVRP.

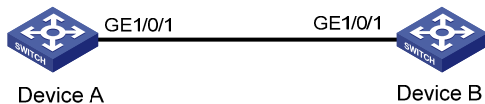
GVRP fixed registration mode configuration example

Network requirements

As shown in [Figure 55](#):

- Device A and Device B are connected through their ports GigabitEthernet 1/0/1.
- Enable GVRP and configure the fixed registration mode on ports to enable the registration and deregistration of static VLAN information between the two devices.

Figure 55 Network diagram



Configuration procedure

1. Configure Device A:

Enable GVRP globally.

```
<DeviceA> system-view
```

```
[DeviceA] gvrp
```

Configure port GigabitEthernet 1/0/1 as a trunk port, and assign it to all VLANs.

```
[DeviceA] interface gigabitethernet 1/0/1
```

```
[DeviceA-GigabitEthernet1/0/1] port link-type trunk
```

```
[DeviceA-GigabitEthernet1/0/1] port trunk permit vlan all
```

Enable GVRP on GigabitEthernet 1/0/1 and set the GVRP registration mode to fixed on the port.

```
[DeviceA-GigabitEthernet1/0/1] gvrp
```

```
[DeviceA-GigabitEthernet1/0/1] gvrp registration fixed
```

```
[DeviceA-GigabitEthernet1/0/1] quit
```

Create VLAN 2 (a static VLAN).

```
[DeviceA] vlan 2
```

```
[DeviceA-vlan2] quit
```

2. Configure Device B:

Enable GVRP globally.

```
<DeviceB> system-view
```

```

[DeviceB] gvrp
# Configure port GigabitEthernet 1/0/1 as a trunk port, and assign it to all VLANs.
[DeviceB] interface gigabitethernet 1/0/1
[DeviceB-GigabitEthernet1/0/1] port link-type trunk
[DeviceB-GigabitEthernet1/0/1] port trunk permit vlan all
# Enable GVRP on GigabitEthernet 1/0/1, and set the GVRP registration mode to fixed on the
port.
[DeviceB-GigabitEthernet1/0/1] gvrp
[DeviceB-GigabitEthernet1/0/1] gvrp registration fixed
[DeviceB-GigabitEthernet1/0/1] quit
# Create VLAN 3 (a static VLAN).
[DeviceB] vlan 3
[DeviceB-vlan3] quit

```

3. Verify the configuration:

Use the **display gvrp local-vlan** command to display the local VLAN information that GVRP maintains on ports. For example:

Display the local VLAN information that GVRP maintains on port GigabitEthernet 1/0/1 of Device A.

```

[DeviceA] display gvrp local-vlan interface gigabitethernet 1/0/1
Following VLANs exist in GVRP local database:
 1(default), 2

```

According to the output, information about VLAN 1 and static VLAN information of VLAN 2 on the local device are registered through GVRP, but dynamic VLAN information of VLAN 3 on Device B is not.

Display the local VLAN information that GVRP maintains on port GigabitEthernet 1/0/1 of Device B.

```

[DeviceB] display gvrp local-vlan interface gigabitethernet 1/0/1
Following VLANs exist in GVRP local database:
 1(default), 3

```

According to the output, information about VLAN 1 and static VLAN information of VLAN 3 on the local device are registered through GVRP, but dynamic VLAN information of VLAN 2 on Device A is not.

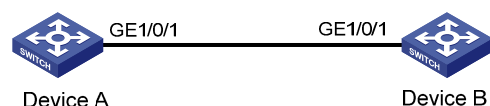
GVRP forbidden registration mode configuration example

Network requirements

As shown in [Figure 56](#):

- Device A and Device B are connected through their ports GigabitEthernet 1/0/1.
- Enable GVRP and configure the forbidden registration mode on ports to prevent the registration and deregistration of all VLANs but VLAN 1 between the two devices.

Figure 56 Network diagram



Configuration procedure

1. Configure Device A:

Enable GVRP globally.

```
<DeviceA> system-view
```

```
[DeviceA] gvrp
```

Configure port GigabitEthernet 1/0/1 as a trunk port, and assign it to all VLANs.

```
[DeviceA] interface gigabitethernet 1/0/1
```

```
[DeviceA-GigabitEthernet1/0/1] port link-type trunk
```

```
[DeviceA-GigabitEthernet1/0/1] port trunk permit vlan all
```

Enable GVRP on GigabitEthernet 1/0/1, and set the GVRP registration mode to forbidden on the port.

```
[DeviceA-GigabitEthernet1/0/1] gvrp
```

```
[DeviceA-GigabitEthernet1/0/1] gvrp registration forbidden
```

```
[DeviceA-GigabitEthernet1/0/1] quit
```

Create VLAN 2 (a static VLAN).

```
[DeviceA] vlan 2
```

```
[DeviceA-vlan2] quit
```

2. Configure Device B:

Enable GVRP globally.

```
<DeviceB> system-view
```

```
[DeviceB] gvrp
```

Configure port GigabitEthernet 1/0/1 as a trunk port, and assign it to all VLANs.

```
[DeviceB] interface gigabitethernet 1/0/1
```

```
[DeviceB-GigabitEthernet1/0/1] port link-type trunk
```

```
[DeviceB-GigabitEthernet1/0/1] port trunk permit vlan all
```

Enable GVRP on GigabitEthernet 1/0/1, and set the GVRP registration mode to forbidden on the port.

```
[DeviceB-GigabitEthernet1/0/1] gvrp
```

```
[DeviceB-GigabitEthernet1/0/1] gvrp registration forbidden
```

```
[DeviceB-GigabitEthernet1/0/1] quit
```

Create VLAN 3 (a static VLAN).

```
[DeviceB] vlan 3
```

```
[DeviceB-vlan3] quit
```

3. Verify the configuration:

Use the **display gvrp local-vlan** command to display the local VLAN information that GVRP maintains on ports. For example:

Display the local VLAN information that GVRP maintains on port GigabitEthernet 1/0/1 of Device A.

```
[DeviceA] display gvrp local-vlan interface gigabitethernet 1/0/1
```

```
Following VLANs exist in GVRP local database:
```

```
1(default)
```

According to the output, information about VLAN 1 is registered through GVRP, but static VLAN information of VLAN 2 on the local device and dynamic VLAN information of VLAN 3 on Device B are not.

Display the local VLAN information that GVRP maintains on port GigabitEthernet 1/0/1 of Device B.

```
[DeviceB] display gvrp local-vlan interface gigabitethernet 1/0/1
```

Following VLANs exist in GVRP local database:

```
1(default)
```

According to the output, information about VLAN 1 is registered through GVRP, but static VLAN information of VLAN 3 on the local device and dynamic VLAN information of VLAN 2 on Device A are not.

Configuring QinQ

Throughout this document, customer network VLANs (CVLANs), also called inner VLANs, refer to the VLANs that a customer uses on the private network; and service provider network VLANs (SVLANs), also called outer VLANs, refer to the VLANs that a service provider uses to carry VLAN tagged traffic for customers.

Overview

802.1Q-in-802.1Q (QinQ) is a flexible, easy-to-implement Layer 2 VPN technology based on IEEE 802.1Q. QinQ enables the edge device on a service provider network to insert an outer VLAN tag in the Ethernet frames from customer networks, so that the Ethernet frames travel across the service provider network (public network) with double VLAN tags. QinQ enables a service provider to use a single SVLAN to serve customers who have multiple CVLANs.

Background and benefits

The IEEE 802.1Q VLAN tag uses 12 bits for VLAN IDs. A device supports a maximum of 4094 VLANs. This is far from enough for isolating users in actual networks, especially in metropolitan area networks (MANs).

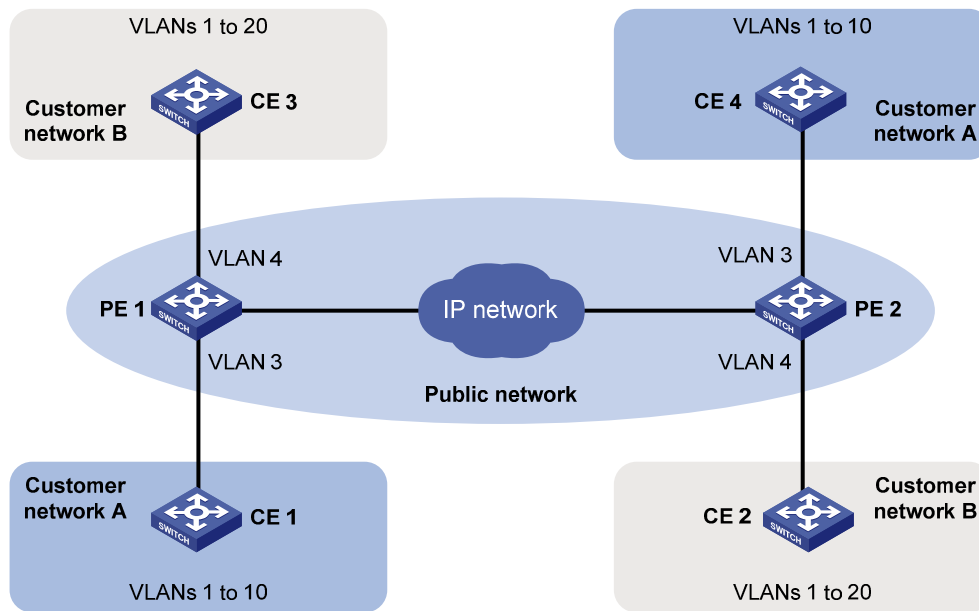
By tagging tagged frames, QinQ expands the available VLAN space from 4094 to 4094×4094 . QinQ delivers the following benefits:

- Releases the stress on the SVLAN resource.
- Enables customers to plan their CVLANs without conflicting with SVLANs.
- Provides an easy-to-implement Layer 2 VPN solution for small-sized MANs or intranets.
- Enables the customers to keep their VLAN assignment schemes unchanged when the service provider upgrades the service provider network.

How QinQ works

The devices in the public network forward a frame only according to its outer VLAN tag and obtain its source MAC address into the MAC address table of the outer VLAN. The inner VLAN tag of the frame is transmitted as the payload.

Figure 57 Typical QinQ application scenario



As shown in [Figure 57](#), customer network A has CVLANs 1 through 10, and customer network B has CVLANs 1 through 20. The service provider assigns SVLAN 3 for customer network A, and assigns SVLAN 4 for customer network B.

When a tagged Ethernet frame from customer network A arrives at a provider edge device (PE), the PE tags the frame with outer VLAN 3. When a tagged Ethernet frame from customer network B arrives at a PE, the PE tags the frame with outer VLAN 4. There is no overlap of VLAN IDs among customers, and traffic from different customers can be identified separately.

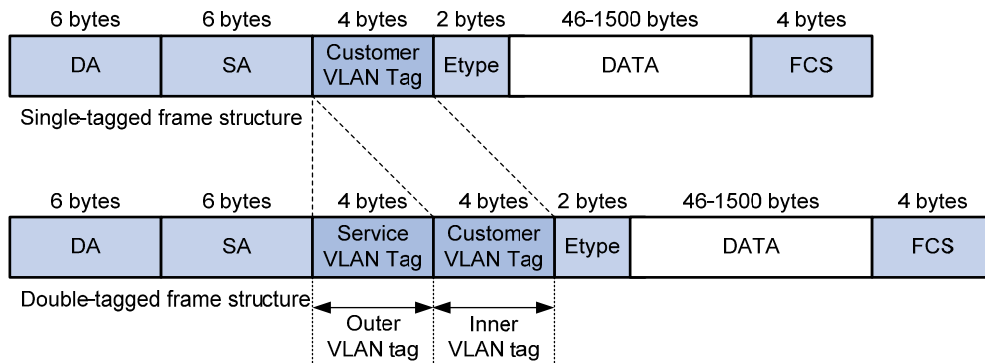
The double-tagged Ethernet frame is then transmitted over the service provider network and arrives at the other PE. The PE removes the SVLAN of the frame before sending it to the target customer edge device (CE).

QinQ frame structure

A QinQ frame is transmitted double-tagged over the service provider network. As shown in [Figure 58](#), the inner VLAN tag is the CVLAN tag, and the outer one is the SVLAN tag that the service provider has allocated to the customer.

QinQ uses CVLAN tags to transmit frames over the private network, and uses SVLAN tags to transmit frames over the public network. When a QinQ frame is transmitted over the public network, its CVLAN tag is transmitted as the payload.

Figure 58 Single-tagged Ethernet frame header and double-tagged Ethernet frame header



The default maximum transmission unit (MTU) of an interface is 1500 bytes. The size of an outer VLAN tag is 4 bytes. HP recommends you to increase the MTU of each interface on the service provider network to at least 1504 bytes.

Implementations of QinQ

HP provides the following QinQ implementations: basic QinQ and selective QinQ.

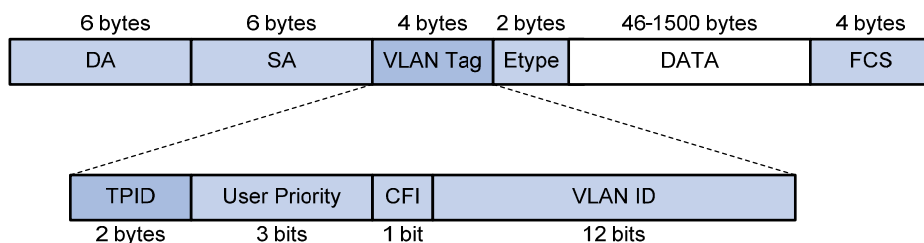
- Basic QinQ
 - Basic QinQ enables a port to tag any incoming frames with its port VLAN ID (PVID) tag, regardless of whether they have been tagged or not. If an incoming frame has been tagged, it becomes a double-tagged frame. If not, it becomes a frame tagged with the PVID tag.
- Selective QinQ
 - Selective QinQ is more flexible than basic QinQ. In addition to all the functions of basic QinQ, selective QinQ enables a port to perform the following per-CVLAN actions for incoming frames:
 - Tag frames from different CVLANs with different SVLAN tags.
 - Mark the outer VLAN 802.1p priority based on the existing inner VLAN 802.1p priority.
 - Modify the inner VLAN ID (available only on the 5500 EI).
 - Besides being able to separate the service provider network from the customer networks, selective QinQ provides abundant service features and enables more flexible networking.

Modifying the TPID in a VLAN tag

A VLAN tag uses the tag protocol identifier (TPID) field to identify the protocol type of the tag. The default value of this field, as defined in IEEE 802.1Q, is 0x8100.

Figure 59 shows the 802.1Q-defined tag structure of an Ethernet frame.

Figure 59 VLAN tag structure of an Ethernet frame



The device determines whether a received frame carries an SVLAN or CVLAN tag by checking the TPID value. For example, if a frame carries an SVLAN tag with TPID value 0x9100 and a CVLAN tag with TPID value 0x8100 and the configured TPID value of the SVLAN tag is 0x9100 and that of the CVLAN tag is 0x8200, the device considers that the frame carries only the SVLAN tag but not the CVLAN tag.

Devices of different vendors may set the TPID of the outer VLAN tag of QinQ frames to different values. For compatibility with these devices, modify the TPID value so that the QinQ frames, when sent to the public network, carry the TPID value identical to the value of a particular vendor, allowing interoperability with the devices of that vendor.

The TPID in an Ethernet frame has the same position as the protocol type field in a frame without a VLAN tag. To avoid problems in packet forwarding and handling in the network, do not set the TPID value to any of the reserved values.

Table 21 Reserved protocol type values

Protocol type	Value
ARP	0x0806
PUP	0x0200
RARP	0x8035
IP	0x0800
IPv6	0x86DD
PPPoE	0x8863/0x8864
MPLS	0x8847/0x8848
IPX/SPX	0x8137
IS-IS	0x8000
LACP	0x8809
802.1X	0x888E
Cluster	0x88A7
Reserved	0xFFFD/0xFFFE/0xFFFF

Protocols and standards

IEEE 802.1Q: *IEEE standard for local and metropolitan area networks: Virtual Bridged Local Area Networks*

QinQ configuration task list

When you configure QinQ, follow these guidelines:

- QinQ requires configurations only on the service provider network.
- QinQ configurations made in Ethernet interface view take effect on the current interface only. Those made in Layer 2 aggregate interface view take effect on the current aggregate interface and all the member ports in the aggregation group. Those made in port group view take effect on all member ports in the current port group.
- Do not configure QinQ on a reflector port. For more information about reflector ports, see *Network Management and Monitoring Configuration Guide*.

- On a port with basic QinQ enabled, you must configure the port to allow packets from its PVID to pass through. On a port with selective QinQ enabled, you must configure the port to allow packets from the inner and outer VLANs of QinQ packets to pass through.

Complete the follows tasks to configure QinQ:

Task	Remarks	
Configuring basic QinQ	Enabling basic QinQ	Required
	Configuring VLAN transparent transmission	Optional
Configuring selective QinQ	Configuring an outer VLAN tagging policy	Required Perform at least one of these tasks.
	Configuring an inner-outer VLAN 802.1p priority mapping	
	Configuring inner VLAN ID substitution (available only on the 5500 E1)	
Configuring the TPID value in VLAN tags	Optional	

Configuring basic QinQ

Enabling basic QinQ

A basic QinQ-enabled port tags an incoming packet with its PVID tag.

To enable basic QinQ:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Enable basic QinQ.	qinq enable	Disabled by default.

Configuring VLAN transparent transmission

When basic QinQ is enabled on a port, all packets passing through the port are tagged with the port's PVID tag. However, by configuring the VLAN transparent transmission function on a port, you can specify the port not to add its PVID tag to packets carrying specific inner VLAN tags when they pass through it, so that these packets are transmitted in the service provider network with single tags.

Configuration restrictions and guidelines

- When you are configuring transparent transmission for a VLAN, you must configure all the devices on the transmission path to permit packets of this VLAN to pass through.

- For VLANs whose packets are to be transparently transmitted through a port, do not configure VLAN mapping for them on the port. For more information about VLAN mapping, see "[Configuring VLAN mapping \(available only on the 5500 EI\)](#)."

Configuration procedure

To configure VLAN transparent transmission:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> • Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> • Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Configure the link type of the ports.	port link-type { hybrid trunk }	N/A
4. Configure the ports to allow packets from the transparent VLANs, and inner and outer VLANs of QinQ packets to pass through.	<ul style="list-style-type: none"> • When the ports are hybrid ports: port hybrid vlan <i>vlan-id-list</i> { tagged untagged } • When the ports are trunk ports: port trunk permit vlan { <i>vlan-id-list</i> all } 	Use either command.
5. Enable basic QinQ on the ports.	qinq enable	By default, basic QinQ is disabled on ports.
6. Configure VLAN transparent transmission on the ports.	qinq transparent-vlan <i>vlan-list</i>	By default, VLAN transparent transmission is not configured.

Configuring selective QinQ

Configuring an outer VLAN tagging policy

Basic QinQ can only tag received frames with the PVID tag of the receiving port. Selective QinQ allows adding different outer VLAN tags based on different inner VLAN tags.

You can configure an outer VLAN tagging policy in one of the following approaches:

- Port-based approach
- QoS policy-based approach (available only on the 5500 EI)

If multiple outer VLAN tagging policies are configured in both approaches on the same port, the one in the QoS policy-based approach takes effect.

Configuring an outer VLAN tagging policy in the port-based approach

Before enabling selective QinQ on a port, enable basic QinQ on the port first. When both features are enabled on the port, frames that meet the selective QinQ condition are handled with selective QinQ on this port first, and the left frames are handled with basic QinQ.

To configure an outer VLAN tagging policy in the port-based approach:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Enable basic QinQ.	qinq enable	Disabled by default.
4. Enter QinQ view and configure the SVLAN tag for the port to add.	qinq vid <i>vlan-id</i>	By default, the SVLAN tag to be added is the PVID tag of the receiving port.
5. Tag frames of the specified CVLANs with the current SVLAN.	raw-vlan-id inbound { all <i>vlan-list</i> }	N/A

NOTE:

- An inner VLAN tag corresponds to only one outer VLAN tag.
- To change an outer VLAN tag, you must delete the old outer VLAN tag configuration and configure a new outer VLAN tag.

Configuring an outer VLAN tagging policy in the QoS policy-based approach (available only on the 5500 EI)

To enable the switch to tag tagged packets based on inner VLAN tags, follow these steps:

- Configure a class to match packets with certain tags.
- Configure a traffic behavior to tag packets with an outer VLAN tag.
- Create a QoS policy and associate the class with the behavior in the policy.
- Apply the QoS policy to the port that connects to the user.

To configure selective QinQ:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a class and enter class view.	traffic classifier <i>classifier-name</i> [operator { and or }]	By default, the operator of a class is AND.
3. Specify the inner VLAN IDs of matching frames.	if-match customer-vlan-id <i>vlan-id-list</i>	N/A
4. Return to system view.	quit	N/A
5. Create a traffic behavior and enter traffic behavior view.	traffic behavior <i>behavior-name</i>	N/A
6. Specify an outer VLAN ID.	nest top-most vlan-id <i>vlan-id</i>	N/A
7. Return to system view.	quit	N/A
8. Create a QoS policy and enter QoS policy view.	qos policy <i>policy-name</i>	N/A

Step	Command	Remarks
9. Associate the traffic class with the traffic behavior defined earlier.	classifier classifier-name behavior behavior-name	N/A
10. Return to system view.	quit	N/A
11. Enter Ethernet interface view or port group view of the customer network-side port.	<ul style="list-style-type: none"> Enter Ethernet interface view: interface interface-type interface-number Enter port group view: port-group manual port-group-name 	Use either command.
12. Enable basic QinQ.	qinq enable	N/A
13. Apply the QoS policy to the incoming traffic.	qos apply policy policy-name inbound	N/A

NOTE:

Selective QinQ has higher priority than basic QinQ. A received frame is tagged with an outer VLAN ID based on basic QinQ only after it fails to match the match criteria defined in the traffic class.

Configuring an inner-outer VLAN 802.1p priority mapping

Through QoS policies, the 5500 SI&5500 EI switches achieve the following inner-outer VLAN 802.1p priority mapping modes:

- Marking the 802.1p priorities in outer VLAN tags according to the inner VLAN IDs or the 802.1p priorities in the inner VLAN tags.
- Copying the 802.1p priority in the inner VLAN tags to the outer VLAN tags.

To mark the 802.1p priorities in outer VLAN tags according to the inner VLAN IDs or the 802.1p priorities in the inner VLAN tags:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a class and enter class view.	traffic classifier classifier-name [operator { and or }]	By default, the operator of a class is AND.
3. Configure a match criterion.	<ul style="list-style-type: none"> Match the specified inner VLAN IDs: if-match customer-vlan-id vlan-id-list Match the specified inner VLAN tag priorities: if-match customer-dot1p 8021p-list 	Use either command.
4. Return to system view.	quit	N/A
5. Create a traffic behavior and enter traffic behavior view.	traffic behavior behavior-name	N/A

Step	Command	Remarks
6. Configure a marking action or an inner-to-outer tag priority copying action.	<ul style="list-style-type: none"> Mark the 802.1p priorities in outer VLAN tags: remark dot1p 8021p Copy the inner 802.1p priorities to outer 802.1p priorities: remark dot1p customer-dot1p-trust 	Use either command.
7. Return to system view.	quit	N/A
8. Create a QoS policy and enter QoS policy view.	qos policy policy-name	N/A
9. Associate the traffic class with the traffic behavior defined earlier.	classifier classifier-name behavior behavior-name	N/A
10. Return to system view.	quit	N/A
11. Enter Ethernet interface view or port group view of the customer network-side port.	<ul style="list-style-type: none"> Enter Ethernet interface view: interface interface-type interface-number Enter port group view: port-group manual port-group-name 	Use either command.
12. Enable basic QinQ.	qinq enable	N/A
13. Apply the QoS policy to the incoming traffic.	qos apply policy policy-name inbound	N/A

Configuring inner VLAN ID substitution (available only on the 5500 EI)

When you configure basic QinQ or an outer VLAN tagging policy for selective QinQ on a port connecting the device to the customer network, basic QinQ or the outer VLAN tagging policy can add outer VLAN tags to packets, but cannot modify the inner VLAN IDs of packets.

To modify the inner VLAN IDs of packets, follow these steps on the port connecting the device to the service provider network:

1. Configure a class to match the specified inner VLAN IDs and outer VLAN ID of packets.
2. Configure a behavior to modify the inner VLAN IDs of packets.
3. Associate the class with the behavior in a QoS policy.
4. Apply the QoS policy to the outgoing traffic of the port connecting the device to the service provider network.

To configure inner VLAN ID substitution:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a class and enter class view.	traffic classifier classifier-name [operator and]	By default, the operator of a class is AND.
3. Configure a match criterion to match the specified inner VLAN IDs.	if-match customer-vlan-id vlan-id-list	N/A

Step	Command	Remarks
4.	Configure a match criterion to match the specified outer VLAN IDs. if-match service-vlan-id <i>vlan-id</i>	N/A
5.	Return to system view. quit	N/A
6.	Create a traffic behavior and enter traffic behavior view. traffic behavior <i>behavior-name</i>	N/A
7.	Configure the action of marking the inner VLAN IDs. remark customer-vlan-id <i>vlan-id</i>	N/A
8.	Return to system view. quit	N/A
9.	Create a QoS policy and enter QoS policy view. qos policy <i>policy-name</i>	N/A
10.	Associate the traffic class with the traffic behavior defined earlier. classifier <i>classifier-name</i> behavior <i>behavior-name</i>	N/A
11.	Return to system view. quit	N/A
12.	Enter Ethernet interface view or port group view of the service provider network-side ports. <ul style="list-style-type: none">• Enter Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i>• Enter port group view: port-group manual <i>port-group-name</i>	Use either command.
13.	Apply the QoS policy to the outgoing traffic. qos apply policy <i>policy-name</i> outbound	N/A

Configuring the TPID value in VLAN tags

Configuring the TPID value on the 5500 EI

To configure the TPID value:

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Configure the TPID value. qinq ethernet-type { customer-tag service-tag } <i>hex-value</i>	Optional. By default, the TPID value is 0x8100. The configuration applies to all ports.

Configuring the TPID value on the 5500 SI

To configure the TPID value:

Step	Command	Remarks
1.	Enter system view. system-view	N/A

Step	Command	Remarks
2. Configure the TPID value.	qinq ethernet-type <i>hex-value</i>	Optional. By default, the TPID value is 0x8100. The configuration applies to all ports.

NOTE:

The TPID value configured on the 5500 SI Switch Series applies to both the CVLAN tags and the SVLAN tags.

QinQ configuration examples

Basic QinQ configuration example

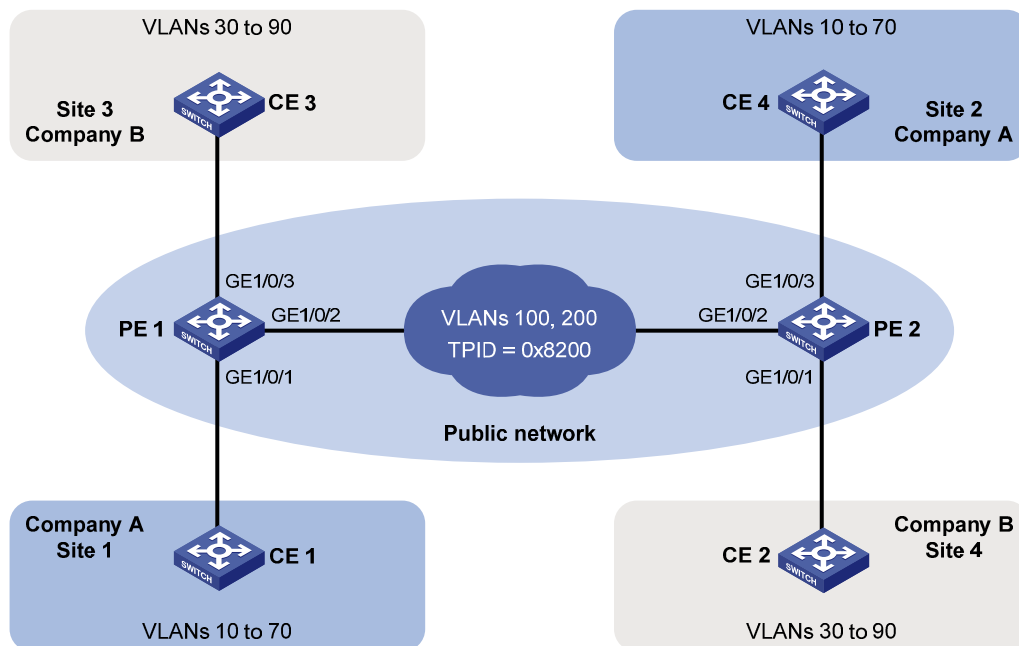
Network requirements

As shown in Figure 60:

- The two branches of Company A, Site 1 and Site 2, are connected through the service provider network and use CVLANs 10 through 70. The two branches of Company B, Site 3 and Site 4, are connected through the service provider network and use CVLANs 30 through 90.
- PE 1 and PE 2 are edge devices on the service provider network and are connected through third-party devices with a TPID value of 0x8200.

Configure the edge and third-party devices to enable communication between the branches of Company A through SVLAN 100, and communication between the branches of Company B through SVLAN 200.

Figure 60 Network diagram



Configuration procedure

! **IMPORTANT:**

Make sure that you have configured the devices in the service provider network to allow QinQ packets to pass through.

1. Configure PE 1:

- Configure GigabitEthernet 1/0/1:

Configure GigabitEthernet 1/0/1 as a trunk port and assign it to VLAN 100 and VLANs 10 through 70..

```
<PE1> system-view
[PE1] interface gigabitethernet 1/0/1
[PE1-GigabitEthernet1/0/1] port link-type trunk
[PE1-GigabitEthernet1/0/1] port trunk permit vlan 100 10 to 70
```

Configure VLAN 100 as the PVID for the port.

```
[PE1-GigabitEthernet1/0/1] port trunk pvid vlan 100
```

Enable basic QinQ on the port.

```
[PE1-GigabitEthernet1/0/1] qinq enable
[PE1-GigabitEthernet1/0/1] quit
```

- Configure GigabitEthernet 1/0/2:

Configure GigabitEthernet 1/0/2 as a trunk port and assign it to VLAN 100 and VLAN 200.

```
[PE1] interface gigabitethernet 1/0/2
[PE1-GigabitEthernet1/0/2] port link-type trunk
[PE1-GigabitEthernet1/0/2] port trunk permit vlan 100 200
```

Set the TPID value in the outer VLAN tag to 0x8200 on the port.

```
[PE1-GigabitEthernet1/0/2] quit
[PE1] qinq ethernet-type service-tag 8200
```

NOTE:

The previous command is available only on the 5500 EI Switch Series. For how to configure the TPID value on the 5500 SI Switch Series, see "[Configuring the TPID value on the 5500 SI.](#)"

- Configure GigabitEthernet 1/0/3:

Configure GigabitEthernet 1/0/3 as a trunk port and assign it to VLAN 200 and VLANs 30 through 90..

```
[PE1] interface gigabitethernet 1/0/3
[PE1-GigabitEthernet1/0/3] port link-type trunk
[PE1-GigabitEthernet1/0/3] port trunk permit vlan 200 30 to 90
```

Configure VLAN 200 as the PVID for the port.

```
[PE1-GigabitEthernet1/0/3] port trunk pvid vlan 200
```

Enable basic QinQ on the port.

```
[PE1-GigabitEthernet1/0/3] qinq enable
[PE1-GigabitEthernet1/0/3] quit
```

2. Configure PE 2:

- Configure GigabitEthernet 1/0/1:

Configure GigabitEthernet 1/0/1 as a trunk port and assign it to VLAN 200 and VLANs 30 through 90..

```
<PE2> system-view
```

- ```
[PE2] interface gigabitethernet 1/0/1
[PE2-GigabitEthernet1/0/1] port link-type trunk
[PE2-GigabitEthernet1/0/1] port trunk permit vlan 200 30 to 90
Configure VLAN 200 as the PVID for the port.
[PE2-GigabitEthernet1/0/1] port trunk pvid vlan 200
Enable basic QinQ on the port.
[PE2-GigabitEthernet1/0/1] qinq enable
[PE2-GigabitEthernet1/0/1] quit
```
- Configure GigabitEthernet 1/0/2:
 

```
Configure GigabitEthernet 1/0/2 as a trunk port and assign it to VLAN 100 and VLAN 200.
[PE2] interface gigabitethernet 1/0/2
[PE2-GigabitEthernet1/0/2] port link-type trunk
[PE2-GigabitEthernet1/0/2] port trunk permit vlan 100 200
Set the TPID value in the outer VLAN tag to 0x8200 on the port.
[PE2-GigabitEthernet1/0/2] quit
[PE2] qinq ethernet-type service-tag 8200
```

---

**NOTE:**

The previous command is available only on the 5500 EI Switch Series. For how to configure the TPID value on the 5500 SI Switch Series, see "[Configuring the TPID value on the 5500 SI.](#)"

---

- Configure GigabitEthernet 1/0/3:
 

```
Configure GigabitEthernet 1/0/3 as a trunk port and assign it to VLAN 100 and VLANs 10 through 70..
[PE2] interface gigabitethernet 1/0/3
[PE2-GigabitEthernet1/0/3] port link-type trunk
[PE2-GigabitEthernet1/0/3] port trunk permit vlan 100 10 to 70
Configure VLAN 100 as the PVID for the port.
[PE2-GigabitEthernet1/0/3] port trunk pvid vlan 100
Enable basic QinQ on the port.
[PE2-GigabitEthernet1/0/3] qinq enable
[PE2-GigabitEthernet1/0/3] quit
```
- 3. Configure third-party devices:**
- On the third-party devices between PE 1 and PE 2, configure the port that connects to PE 1 and that connecting to PE 2 to allow tagged frames of VLAN 100 and VLAN 200 to pass through.

## Selective QinQ configuration example (configuring an outer VLAN tagging policy in port-based approach)

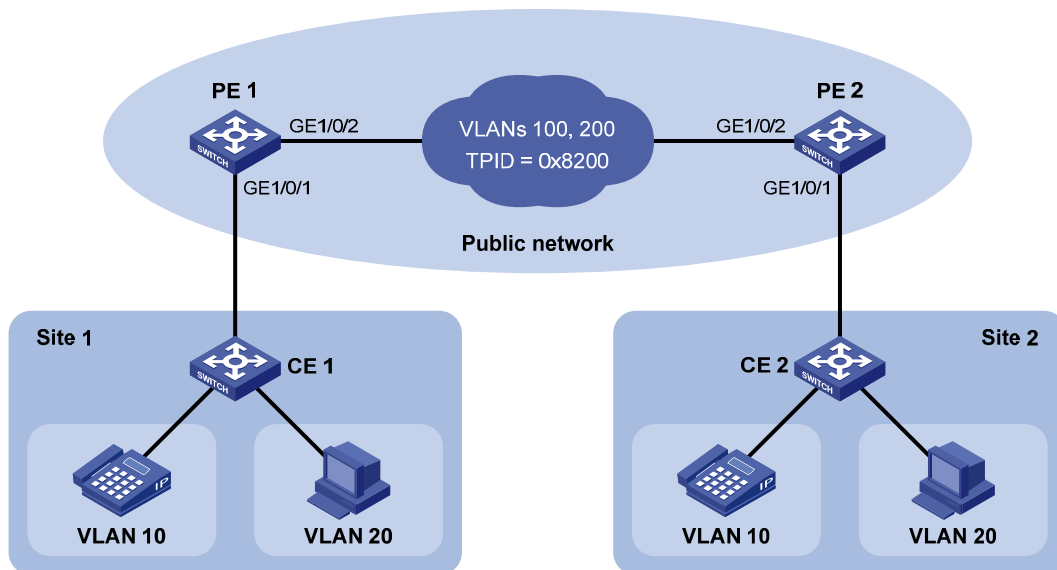
### Network requirements

As shown in [Figure 61](#):

- The two branches of a company, Site 1 and Site 2, are connected through the service provider network and use CVLAN 10 and CVLAN 20 to transmit voice traffic and data traffic separately.
- PE 1 and PE 2 are edge devices on the service provider network and are connected through third-party devices with a TPID value of 0x8200.

Configure the edge and third-party devices to allow frames from CVLAN 10 to be transmitted between the branches via SVLAN 100 and frames from CVLAN 20 to be transmitted between the branches via SVLAN 200.

**Figure 61 Network diagram**



## Configuration procedure

### ! IMPORTANT:

Make sure that you have configured the devices in the service provider network to allow QinQ packets to pass through.

#### 1. Configure PE 1:

- Configure GigabitEthernet 1/0/1:

# Configure GigabitEthernet 1/0/1 as a hybrid port to permit frames of VLAN 10 and VLAN 20 to pass through tagged, and frames of VLAN 100 and VLAN 200 to pass through untagged.

```
<PE1> system-view
[PE1] interface gigabitethernet 1/0/1
[PE1-GigabitEthernet1/0/1] port link-type hybrid
[PE1-GigabitEthernet1/0/1] port hybrid vlan 10 20 tagged
```

```
[PE1-GigabitEthernet1/0/1] port hybrid vlan 100 200 untagged
```

# Enable basic QinQ on the port.

```
[PE1-GigabitEthernet1/0/1] qinq enable
```

# Configure the port to tag VLAN 10 frames with outer VLAN ID 100.

```
[PE1-GigabitEthernet1/0/1] qinq vid 100
[PE1-GigabitEthernet1/0/1-vid-100] raw-vlan-id inbound 10
[PE1-GigabitEthernet1/0/1-vid-100] quit
```

# Configure the port to tag VLAN 20 frames with outer VLAN ID 200.

```
[PE1-GigabitEthernet1/0/1] qinq vid 200
[PE1-GigabitEthernet1/0/1-vid-200] raw-vlan-id inbound 20
```



- ```
[PE1-GigabitEthernet1/0/1-vid-200] quit
[PE1-GigabitEthernet1/0/1] quit
```
- **Configure GigabitEthernet 1/0/2:**

Configure GigabitEthernet 1/0/2 as a trunk port and assign it to VLAN 100 and VLAN 200.

```
[PE1] interface gigabitethernet 1/0/2
[PE1-GigabitEthernet1/0/2] port link-type trunk
[PE1-GigabitEthernet1/0/2] port trunk permit vlan 100 200
```

Set the TPID in the outer VLAN tags to 0x8200.

```
[PE1-GigabitEthernet1/0/2] quit
[PE1] qinq ethernet-type service-tag 8200
```

NOTE:

The previous command is available only on the 5500 EI Switch Series. For how to configure the TPID value on the 5500 SI Switch Series, see "[Configuring the TPID value on the 5500 SI.](#)"

2. Configure PE 2:

- **Configure GigabitEthernet 1/0/1:**

Configure GigabitEthernet 1/0/1 as a hybrid port to permit frames of VLAN 10 and VLAN 20 to pass through tagged, and frames of VLAN 100 and VLAN 200 to pass through untagged.

```
[PE2] interface gigabitethernet 1/0/1
[PE2-GigabitEthernet1/0/1] port link-type hybrid
[PE2-GigabitEthernet1/0/1] port hybrid vlan 10 20 tagged
[PE2-GigabitEthernet1/0/1] port hybrid vlan 100 200 untagged
```

Enable basic QinQ on the port.

```
[PE2-GigabitEthernet1/0/1] qinq enable
```

Configure the port to tag VLAN 10 frames with outer VLAN ID 100.

```
[PE2-GigabitEthernet1/0/1] qinq vid 100
[PE2-GigabitEthernet1/0/1-vid-100] raw-vlan-id inbound 10
[PE2-GigabitEthernet1/0/1-vid-100] quit
```

Configure the port to tag VLAN 20 frames with outer VLAN ID 200.

```
[PE2-GigabitEthernet1/0/1] qinq vid 200
[PE2-GigabitEthernet1/0/1-vid-200] raw-vlan-id inbound 20
[PE2-GigabitEthernet1/0/1-vid-200] quit
[PE2-GigabitEthernet1/0/1] quit
```
- **Configure GigabitEthernet 1/0/2:**

Configure GigabitEthernet 1/0/2 as a trunk port and assign it to VLAN 100 and VLAN 200.

```
[PE2] interface gigabitethernet 1/0/2
[PE2-GigabitEthernet1/0/2] port link-type trunk
[PE2-GigabitEthernet1/0/2] port trunk permit vlan 100 200
```

Set the TPID in the outer VLAN tags to 0x8200.

```
[PE2-GigabitEthernet1/0/2] quit
[PE2] qinq ethernet-type service-tag 8200
```

NOTE:

The previous command is available only on the 5500 EI Switch Series. For how to configure the TPID value on the 5500 SI Switch Series, see "[Configuring the TPID value on the 5500 SI.](#)"

3. Configure third-party devices:

On the third-party devices between PE 1 and PE 2, configure the port that connects to PE 1 and that connecting to PE 2 to allow tagged frames of VLAN 100 and VLAN 200 to pass through.

Selective QinQ configuration example (configuring an outer VLAN tagging policy in QoS policy-based approach)

! IMPORTANT:

This configuration example is available only on the 5500 EI Switch Series.

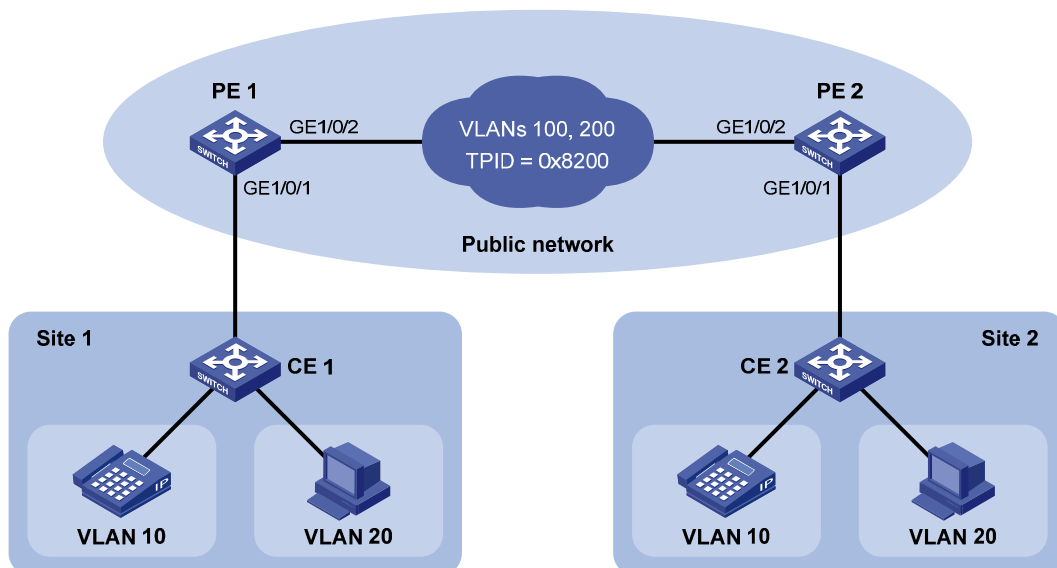
Network requirements

As shown in [Figure 61](#):

- The two branches of a company, Site 1 and Site 2, are connected through the service provider network and use CVLAN 10 and CVLAN 20 to transmit voice traffic and data traffic separately.
- PE 1 and PE 2 are edge devices on the service provider network and are connected through third-party devices with a TPID value of 0x8200.

Configure the edge and third-party devices to allow frames from CVLAN 10 to be transmitted between the branches via SVLAN 100 and frames from CVLAN 20 to be transmitted between the branches via SVLAN 200.

Figure 62 Network diagram



Configuration procedure

❗ **IMPORTANT:**

Make sure that you have configured the devices in the service provider network to allow QinQ packets to pass through.

1. Configure PE 1:

- Configure GigabitEthernet 1/0/1:

Configure GigabitEthernet 1/0/1 as a hybrid port to permit frames of VLAN 10 and VLAN 20 to pass through tagged, and frames of VLAN 100 and VLAN 200 to pass through untagged.

```
<PE1> system-view
[PE1] interface gigabitethernet 1/0/1
[PE1-GigabitEthernet1/0/1] port link-type hybrid
[PE1-GigabitEthernet1/0/1] port hybrid vlan 10 20 tagged
[PE1-GigabitEthernet1/0/1] port hybrid vlan 100 200 untagged
[PE1-GigabitEthernet1/0/1] quit
```

Create class **A10** and configure the class to match frames from Site 1 with CVLAN 10.

```
[PE1] traffic classifier A10
[PE1-classifier-A10] if-match customer-vlan-id 10
[PE1-classifier-A10] quit
```

Create traffic behavior **P100** and add the action of inserting outer VLAN tag 100.

```
[PE1] traffic behavior P100
[PE1-behavior-P100] nest top-most vlan-id 100
[PE1-behavior-P100] quit
```

Create class **A20** and configure the class to match frames with CVLAN 20. Create traffic behavior **P200** and add the action of inserting outer VLAN tag 200.

```
[PE1] traffic classifier A20
[PE1-classifier-A20] if-match customer-vlan-id 20
[PE1-classifier-A20] quit
[PE1] traffic behavior P200
[PE1-behavior-P200] nest top-most vlan-id 200
[PE1-behavior-P200] quit
```

Create a QoS policy named **qinq**, associate traffic class **A10** with traffic behavior **P100**, and associate traffic class **A20** with traffic behavior **P200**.

```
[PE1] qos policy qinq
[PE1-qospolicy-qinq] classifier A10 behavior P100
[PE1-qospolicy-qinq] classifier A20 behavior P200
[PE1-qospolicy-qinq] quit
```

Enable basic QinQ on the port.

```
[PE1] interface gigabitethernet 1/0/1
[PE1-GigabitEthernet1/0/1] qinq enable
```

Apply QoS policy **qinq** to the incoming traffic on the port.

```
[PE1-GigabitEthernet1/0/1] qos apply policy qinq inbound
```

- Configure GigabitEthernet 1/0/2:

Configure GigabitEthernet 1/0/2 as a trunk port and assign it to VLAN 100 and VLAN 200.

```
[PE1] interface gigabitethernet 1/0/2
[PE1-GigabitEthernet1/0/2] port link-type trunk
```

```
[PE1-GigabitEthernet1/0/2] port trunk permit vlan 100 200
# Set the TPID in the outer VLAN tags to 0x8200.
[PE1-GigabitEthernet1/0/2] quit
[PE1] qinq ethernet-type service-tag 8200
```

2. Configure PE 2:

- Configure GigabitEthernet 1/0/1:

Configure GigabitEthernet 1/0/1 as a hybrid port to permit frames of VLAN 10 and VLAN 20 to pass through tagged, and frames of VLAN 100 and VLAN 200 to pass through untagged.

```
[PE2] interface gigabitethernet 1/0/1
[PE2-GigabitEthernet1/0/1] port link-type hybrid
[PE2-GigabitEthernet1/0/1] port hybrid vlan 10 20 tagged
[PE2-GigabitEthernet1/0/1] port hybrid vlan 100 200 untagged
[PE2-GigabitEthernet1/0/1] quit
```

Create class **A10** and configure the class to match frames from Site 2 with CVLAN 10.

```
[PE2] traffic classifier A10
[PE2-classifier-A10] if-match customer-vlan-id 10
[PE2-classifier-A10] quit
```

Create traffic behavior **P100** and add the action of inserting outer VLAN tag 100.

```
[PE2] traffic behavior P100
[PE2-behavior-P100] nest top-most vlan-id 100
[PE2-behavior-P100] quit
```

Create class **A20** and configure the class to match frames with CVLAN 20. Create traffic behavior **P200** and add the action of inserting outer VLAN tag 200.

```
[PE2] traffic classifier A20
[PE2-classifier-A20] if-match customer-vlan-id 20
[PE2-classifier-A20] quit
[PE2] traffic behavior P200
[PE2-behavior-P200] nest top-most vlan-id 200
[PE2-behavior-P200] quit
```

Create a QoS policy named **qinq**, associate traffic class **A10** with traffic behavior **P100**, and associate traffic class **A20** with traffic behavior **P200**.

```
[PE2] qos policy qinq
[PE2-qospolicy-qinq] classifier A10 behavior P100
[PE2-qospolicy-qinq] classifier A20 behavior P200
[PE2-qospolicy-qinq] quit
```

Enable basic QinQ on the port.

```
[PE2] interface gigabitethernet 1/0/1
[PE2-GigabitEthernet1/0/1] qinq enable
```

Apply QoS policy **qinq** to the incoming traffic on the port.

```
[PE2-GigabitEthernet1/0/1] qos apply policy qinq inbound
```

- Configure GigabitEthernet 1/0/2:

Configure GigabitEthernet 1/0/2 as a trunk port and assign it to VLAN 100 and VLAN 200.

```
[PE2] interface gigabitethernet 1/0/2
[PE2-GigabitEthernet1/0/2] port link-type trunk
[PE2-GigabitEthernet1/0/2] port trunk permit vlan 100 200
```

```
# Set the TPID in the outer VLAN tags to 0x8200.
```

```
[PE2-GigabitEthernet1/0/2] quit
```

```
[PE2] qinq ethernet-type service-tag 8200
```

3. Configure third-party devices:

On the third-party devices between PE 1 and PE 2, configure the port that connects to PE 1 and the port that connects to PE 2 to allow tagged frames of VLAN 100 and VLAN 200 to pass through.

Selective QinQ configuration example (inner VLAN ID substitution)

⚠ **IMPORTANT:**

This configuration example is available only on the 5500 EI Switch Series.

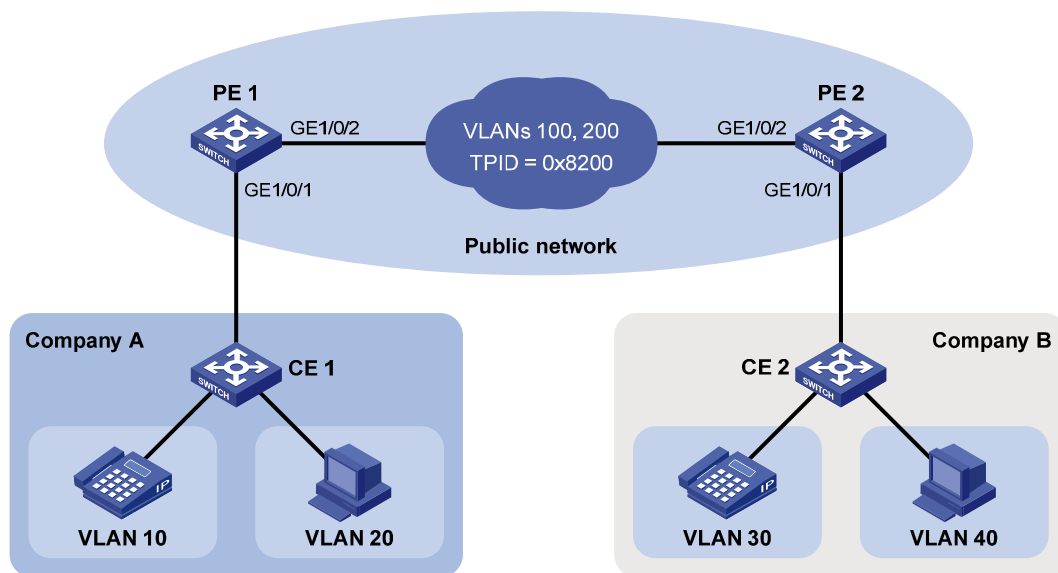
Network requirements

As shown in Figure 63:

- Company A uses CVLAN 10 and CVLAN 20 to transmit voice traffic and data traffic separately, and Company B uses CVLAN 30 and CVLAN 40 to transmit voice traffic and data traffic separately. The services of the two companies are integrated together.
- PE 1 and PE 2 are edge devices on the service provider network and are connected through third-party devices with a TPID value of 0x8200.

Configure the edge and third-party devices to allow the voice traffic and data traffic to be transmitted between the two companies via SVLAN 100 SVLAN 200 separately.

Figure 63 Network diagram



Configuration procedure

⚠ **IMPORTANT:**

Make sure that you have configured the devices in the service provider network to allow QinQ packets to pass through.

1. Configure PE 1:

o Configure GigabitEthernet 1/0/1:

Configure GigabitEthernet 1/0/1 as a hybrid port to permit frames of VLAN 10 and VLAN 20 to pass through tagged, and frames of VLAN 100 and VLAN 200 to pass through untagged.

```
<PE1> system-view
[PE1] interface gigabitethernet 1/0/1
[PE1-GigabitEthernet1/0/1] port link-type hybrid
[PE1-GigabitEthernet1/0/1] port hybrid vlan 10 20 tagged
[PE1-GigabitEthernet1/0/1] port hybrid vlan 100 200 untagged
[PE1-GigabitEthernet1/0/1] quit
```

Create class **A10** and configure the class to match frames with CVLAN 10.

```
[PE1] traffic classifier A10
[PE1-classifier-A10] if-match customer-vlan-id 10
[PE1-classifier-A10] quit
```

Create traffic behavior **P100** and add the action of inserting outer VLAN tag 100.

```
[PE1] traffic behavior P100
[PE1-behavior-P100] nest top-most vlan-id 100
[PE1-behavior-P100] quit
```

Create class **A20** and configure the class to match frames with CVLAN 20. Create traffic behavior **P200** and add the action of inserting outer VLAN tag 200.

```
[PE1] traffic classifier A20
[PE1-classifier-A20] if-match customer-vlan-id 20
[PE1-classifier-A20] quit
[PE1] traffic behavior P200
[PE1-behavior-P200] nest top-most vlan-id 200
[PE1-behavior-P200] quit
```

Create a QoS policy named **qinq**, associate traffic class **A10** with traffic behavior **P100**, and associate traffic class **A20** with traffic behavior **P200**.

```
[PE1] qos policy qinq
[PE1-qospolicy-qinq] classifier A10 behavior P100
[PE1-qospolicy-qinq] classifier A20 behavior P200
[PE1-qospolicy-qinq] quit
```

Enable basic QinQ on the port.

```
[PE1] interface gigabitethernet 1/0/1
[PE1-GigabitEthernet1/0/1] qinq enable
```

Apply QoS policy **qinq** to the incoming traffic on the port.

```
[PE1-GigabitEthernet1/0/1] qos apply policy qinq inbound
```

o Configure GigabitEthernet 1/0/2:

Configure GigabitEthernet 1/0/2 as a trunk port and assign it to VLAN 100 and VLAN 200.

```
[PE1] interface gigabitethernet 1/0/2
[PE1-GigabitEthernet1/0/2] port link-type trunk
[PE1-GigabitEthernet1/0/2] port trunk permit vlan 100 200
```

Create class **A100** and configure the class to match frames with CVLAN 10 and SVLAN 100.

```

[PE1] traffic classifier A100
[PE1-classifier-A100] if-match customer-vlan-id 10
[PE1-classifier-A100] if-match service-vlan-id 100
[PE1-classifier-A100] quit
# Configure traffic behavior T100 to mark matching packets with CVLAN 30.
[PE1] traffic behavior T100
[PE1-behavior-T100] remark customer-vlan-id 30
[PE1-behavior-T100] quit
# Create class A200 and configure the class to match frames with CVLAN 20 and SVLAN 200.
[PE1] traffic classifier A200
[PE1-classifier-A200] if-match customer-vlan-id 20
[PE1-classifier-A200] if-match service-vlan-id 200
[PE1-classifier-A200] quit
# Configure traffic behavior T200 to mark matching packets with CVLAN 40.
[PE1] traffic behavior T200
[PE1-behavior-T200] remark customer-vlan-id 40
[PE1-behavior-T200] quit
# Create a QoS policy named sqinq, associate traffic class A100 with traffic behavior T100, and associate traffic class A200 with traffic behavior T200.
[PE1] qos policy sqinq
[PE1-qospolicy-sqinq] classifier A100 behavior T100
[PE1-qospolicy-sqinq] classifier A200 behavior T200
[PE1-qospolicy-sqinq] quit
# Apply QoS policy sqinq to the outgoing traffic on the port.
[PE1-GigabitEthernet1/0/2] qos apply policy sqinq outbound
# Set the TPID value in the outer tag to 0x8200.
[PE1-GigabitEthernet1/0/2] quit
[PE1] qinq ethernet-type service-tag 8200

```

2. Configure PE 2:

- Configure GigabitEthernet 1/0/1:

```

# Configure GigabitEthernet 1/0/1 as a hybrid port to permit frames of VLAN 30 and VLAN 40 to pass through tagged, and frames of VLAN 100 and VLAN 200 to pass through untagged.
[PE2] interface gigabitethernet 1/0/1
[PE2-GigabitEthernet1/0/1] port link-type hybrid
[PE2-GigabitEthernet1/0/1] port hybrid vlan 30 40 tagged
[PE2-GigabitEthernet1/0/1] port hybrid vlan 100 200 untagged
[PE2-GigabitEthernet1/0/1] quit
# Create class A30 and configure the class to match frames with CVLAN 30.
[PE2] traffic classifier A30
[PE2-classifier-A30] if-match customer-vlan-id 30
[PE2-classifier-A30] quit
# Create traffic behavior P100 and add the action of inserting outer VLAN tag 100.
[PE2] traffic behavior P100
[PE2-behavior-P100] nest top-most vlan-id 100

```

```

[PE2-behavior-P100] quit
# Create class A40 and configure the class to match frames with CVLAN 40. Create traffic
behavior P200 and add the action of inserting outer VLAN tag 200.
[PE2] traffic classifier A40
[PE2-classifier-A40] if-match customer-vlan-id 40
[PE2-classifier-A40] quit
[PE2] traffic behavior P200
[PE2-behavior-P200] nest top-most vlan-id 200
[PE2-behavior-P200] quit
# Create a QoS policy named qinq, associate traffic class A30 with traffic behavior P100, and
associate traffic class A40 with traffic behavior P200.
[PE2] qos policy qinq
[PE2-qospolicy-qinq] classifier A30 behavior P100
[PE2-qospolicy-qinq] classifier A40 behavior P200
[PE2-qospolicy-qinq] quit
# Enable basic QinQ on the port.
[PE2] interface gigabitethernet 1/0/1
[PE2-GigabitEthernet1/0/1] qinq enable
# Apply QoS policy qinq to the incoming traffic on the port.
[PE2-GigabitEthernet1/0/1] qos apply policy qinq inbound
o Configure GigabitEthernet 1/0/2:
# Configure GigabitEthernet 1/0/2 as a trunk port and assign it to VLAN 100 and VLAN 200.
[PE2] interface gigabitethernet 1/0/2
[PE2-GigabitEthernet1/0/2] port link-type trunk
[PE2-GigabitEthernet1/0/2] port trunk permit vlan 100 200
# Create class A100 and configure the class to match frames with CVLAN 30 and SVLAN
100.
[PE2] traffic classifier A100
[PE2-classifier-A100] if-match customer-vlan-id 30
[PE2-classifier-A100] if-match service-vlan-id 100
[PE2-classifier-A100] quit
# Configure traffic behavior T100 to mark matching packets with CVLAN 10.
[PE2] traffic behavior T100
[PE2-behavior-T100] remark customer-vlan-id 10
[PE2-behavior-T100] quit
# Create class A200 and configure the class to match frames with CVLAN 40 and SVLAN
200.
[PE2] traffic classifier A200
[PE2-classifier-A200] if-match customer-vlan-id 40
[PE2-classifier-A200] if-match service-vlan-id 200
[PE2-classifier-A200] quit
# Configure traffic behavior T200 to mark matching packets with CVLAN 20.
[PE2] traffic behavior T200
[PE2-behavior-T200] remark customer-vlan-id 20
[PE2-behavior-T200] quit

```


Create a QoS policy named **sqinq**, associate traffic class **A100** with traffic behavior **T100**, and associate traffic class **A200** with traffic behavior **T200**.

```
[PE2] qos policy sqinq
[PE2-qospolicy-sqinq] classifier A100 behavior T100
[PE2-qospolicy-sqinq] classifier A200 behavior T200
[PE2-qospolicy-sqinq] quit
```

Apply QoS policy **sqinq** to the outgoing traffic on the port.

```
[PE2] interface gigabitethernet 1/0/2
[PE1-GigabitEthernet1/0/2] qos apply policy sqinq outbound
```

Set the TPID value in the outer tag to 0x8200.

```
[PE2-GigabitEthernet1/0/2] quit
[PE2] qinq ethernet-type service-tag 8200
```

3. Configure third-party devices:

On the third-party devices between PE 1 and PE 2, configure the port that connects to PE 1 and that connecting to PE 2 to allow tagged frames of VLAN 100 and VLAN 200 to pass through.

Configuring VLAN mapping (available only on the 5500 EI)

Overview

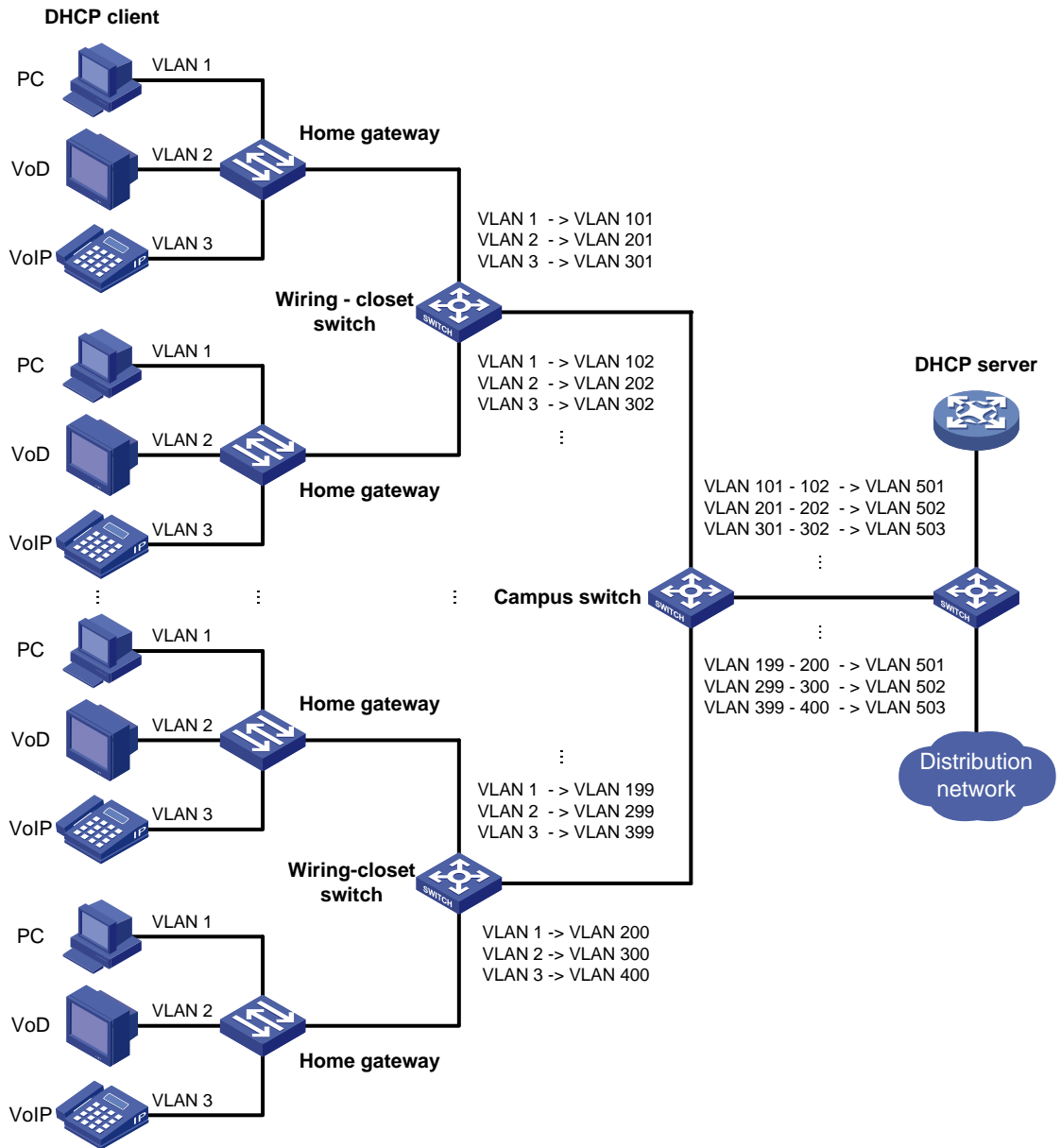
VLAN mapping re-marks VLAN tagged traffic with new VLAN IDs. HP provides the following types of VLAN mapping:

- **One-to-one VLAN mapping**—Replaces one VLAN tag with another. You can use one-to-one VLAN mapping to sub-classify traffic from a particular VLAN for granular QoS control.
- **Many-to-one VLAN mapping**—Replaces multiple VLAN tags with the same VLAN tag. You can use many-to-one VLAN mapping to aggregate traffic from different VLANs to regulate the aggregate traffic as a whole. Many-to-one VLAN mapping is usually used together with one-to-one VLAN mapping.
- **Two-to-two VLAN mapping**—Replaces the outer and inner VLAN IDs of double tagged traffic with a new pair of VLAN IDs. You can use two-to-two VLAN mapping to enable two remote sites in different VLANs to communicate at Layer 2 across two service provider networks that use different VLAN assignment schemes.

Application scenario of one-to-one and many-to-one VLAN mapping

Figure 64 shows a typical application scenario in which each home gateway uses different VLANs to transmit the PC, VoD, and VoIP services.

Figure 64 Application scenario of one-to-one and many-to-one VLAN mapping

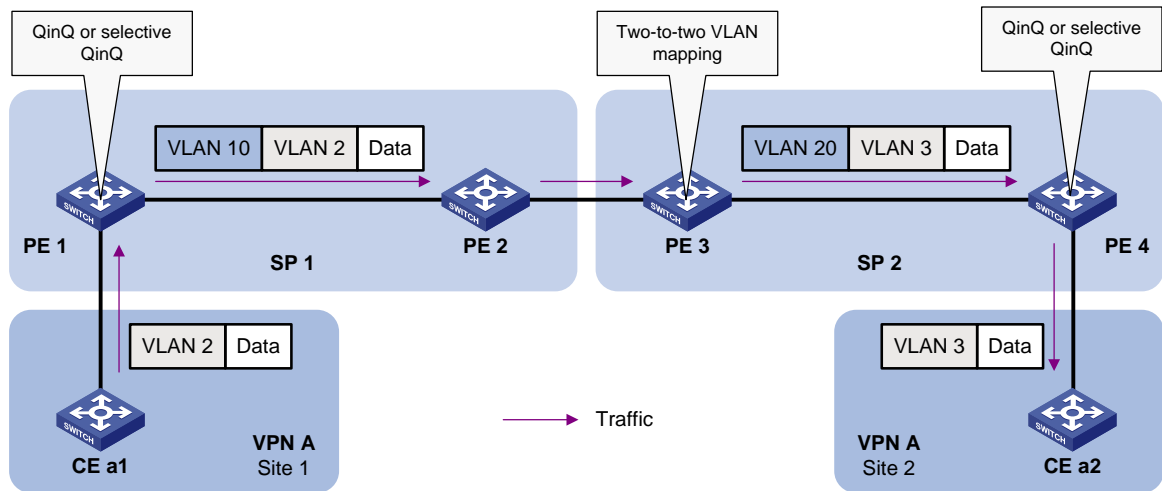


To further sub-classify each type of traffic by customer, perform one-to-one VLAN mapping on the wiring-closet switches, and assign a separate VLAN for each type of traffic from each customer. The required total number of VLANs in the network can be large. To prevent the maximum number of VLANs from being exceeded on the distribution layer device, perform many-to-one VLAN mapping on the campus switch to assign the same type of traffic from different customers to the same VLAN.

Application scenario of two-to-two VLAN mapping

Figure 65 shows a typical application scenario in which two remote sites of VPN A, Site 1 and Site 2, must communicate across two SP networks, SP 1 and SP 2.

Figure 65 Application scenario of two-to-two VLAN mapping



Site 1 and Site 2 are in VLAN 2 and VLAN 3, respectively. The VLAN assigned for VPN A is VLAN 10 in the SP 1 network and VLAN 20 in the SP 2 network.

If Site 1 sends a packet to Site 2, the packet is processed on the way to its destination using the following workflow:

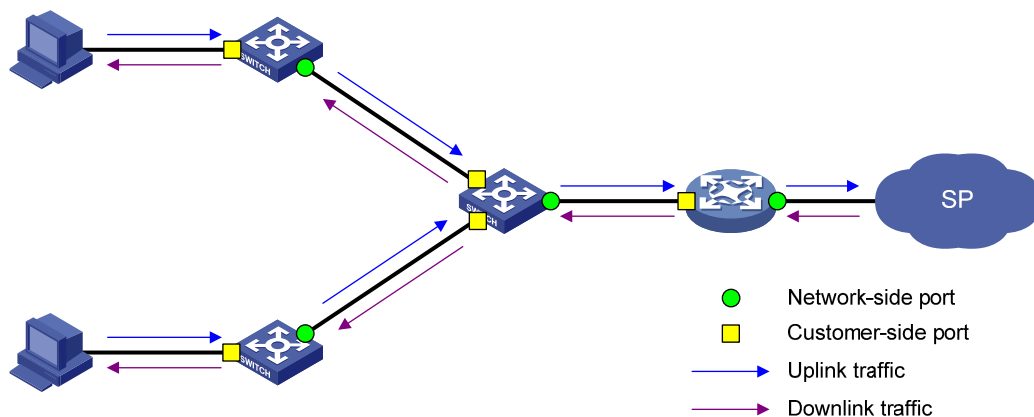
1. When the packet tagged with VLAN 2 arrives at the edge of network SP 1, PE 1 tags the packet with outer VLAN 10 through basic QinQ or selective QinQ.
2. When the double-tagged packet enters the SP 2 network, PE 3 replaces the outer VLAN tag (VLAN 10) with VLAN 20. Because the packet is destined for Site 2 in VLAN 3, PE 3 also replaces the inner tag (VLAN 2) of the packet with VLAN 3. This process is two-to-two VLAN mapping.
3. When PE 4 receives the packet with the new VLAN tag pair, it removes the outer VLAN tag and forwards the packet to VLAN 3.

For more information about basic QinQ and selection QinQ configurations, see "[Configuring QinQ.](#)"

Concepts and terms

Figure 66 shows a simplified network to help explain the concepts and terms that you might encounter when you work with VLAN mapping.

Figure 66 Basic concepts of VLAN mapping



- **Uplink traffic**—Traffic transmitted from the customer network to the service provider network.

- **Downlink traffic**—Traffic transmitted from the service provider network to the customer network.
- **Network-side port**—A port connected to or closer to the service provider network.
- **Customer-side port**—A port connected to or closer to the customer network.
- **Uplink policy**—A QoS policy that defines VLAN mapping rules for uplink traffic.
- **Downlink policy**—A QoS policy that defines VLAN mapping rules for downlink traffic.
- **Customer VLANs (CVLANs)**—VLANs assigned for customers.
- **Service provider VLANs (SVLANs)**—VLANs assigned for transmitting traffic across the service provider network.

For more information about QoS policies, see *ACL and QoS Configuration Guide*.

VLAN mapping implementations

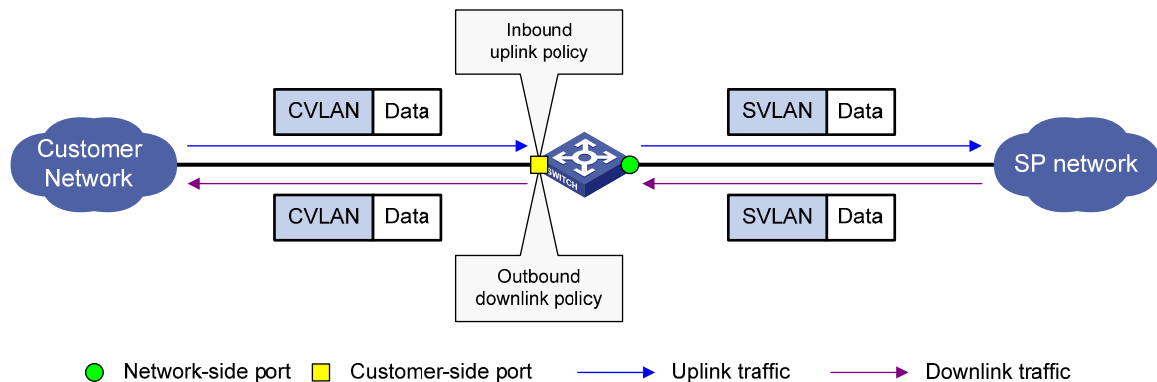
This section describes how VLAN mapping is implemented on your device.

One-to-one VLAN mapping

Implement one-to-one VLAN mapping on the customer-side port through the following configurations, as shown in [Figure 67](#):

- Apply an uplink policy to the incoming traffic, mapping each CVLAN ID to a unique SVLAN ID. When a packet arrives, the switch replaces its CVLAN ID with the matching SVLAN ID.
- Apply a downlink policy to the outgoing traffic, mapping each SVLAN ID back to its corresponding CVLAN ID. When forwarding a packet out of the port, the switch replaces its SVLAN ID with the matching CVLAN ID.

Figure 67 One-to-one VLAN mapping implementation

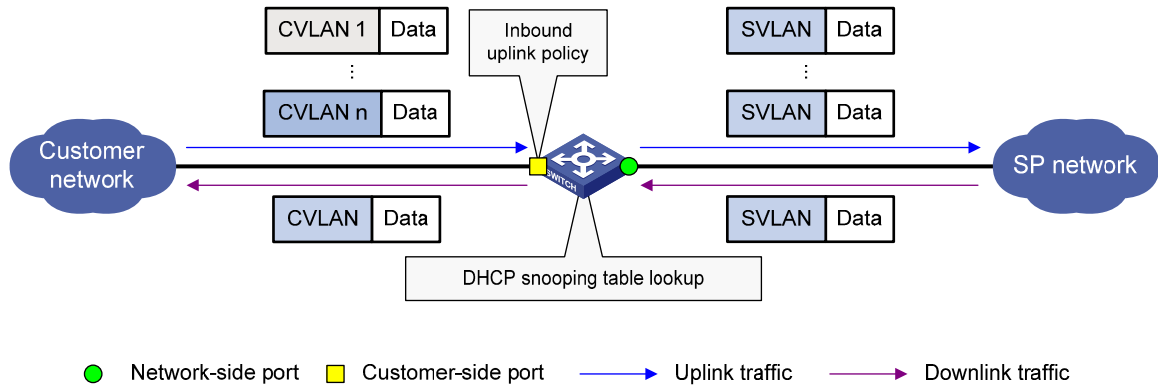


Many-to-one VLAN mapping

Implement many-to-one VLAN mapping through the following configurations, as shown in [Figure 68](#):

- Apply an uplink policy to the incoming traffic on the customer-side port to map different CVLAN IDs to one SVLAN ID. When a packet arrives, the switch replaces its CVLAN tag with the matching SVLAN tag.
- Configure the network-side port as a DHCP snooping trusted port. For downlink traffic, the switch looks through the DHCP snooping table, and replaces the SVLAN ID with the CVLAN ID found in the table.

Figure 68 Many-to-one VLAN mapping implementation



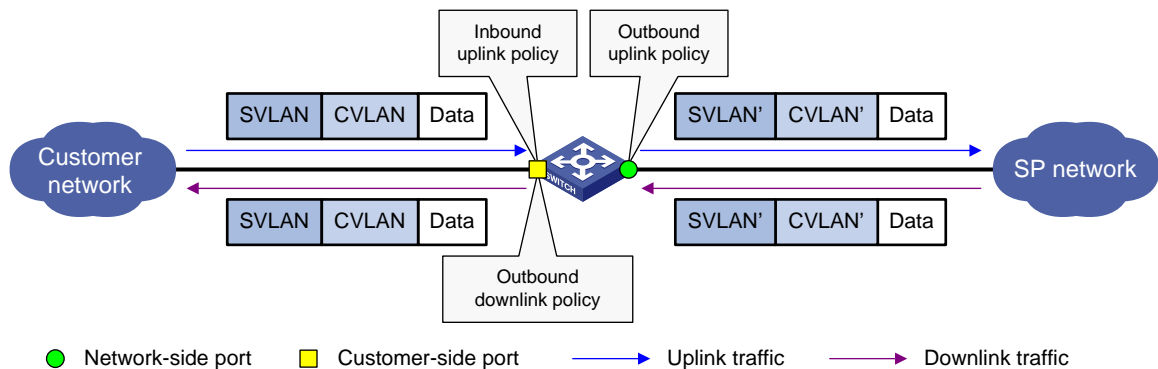
Each DHCP snooping entry contains information about one DHCP client, including its IP address, MAC address, and CVLAN. For more information about DHCP snooping, see *Layer 3—IP Services Configuration Guide*.

Two-to-two VLAN mapping

Implement two-to-two VLAN mapping through the following configurations, as shown in [Figure 69](#).

- For uplink traffic, apply an inbound policy on the customer-side port to replace the SVLAN with a new CVLAN, and apply an outbound policy on the network-side port to replace the CVLAN with a new SVLAN.
- For downlink traffic, apply an outbound policy on the customer-side port to replace the double tags with the original VLAN tag pair.

Figure 69 Two-to-two VLAN mapping implementation



VLAN mapping configuration tasks

Use the VLAN mapping methods as appropriate to the roles of your switches in the network, as described in this table:

Task	Switch role
Configuring one-to-one VLAN mapping	Wiring-closet switch
Configuring many-to-one VLAN mapping	Campus switch
Configuring two-to-two VLAN mapping	Edge switch between SP networks

Configuring one-to-one VLAN mapping

Perform one-to-one VLAN mapping on wiring-closet switches (see [Figure 64](#)) to isolate traffic by both user and traffic type.

Perform these tasks to configure one-to-one VLAN mapping:

Task	Description
Configuring an uplink policy	Creates CVLAN-to-SVLAN mappings (required).
Configuring a downlink policy	Creates SVLAN-to-CVLAN mappings (required).
Configuring the customer-side port	Configures settings required for one-to-one VLAN mapping (required).
Configuring the network-side port	Configures VLAN settings required for normal communication (required).

Configuration prerequisites

Create CVLANs and SVLANs, and plan CVLAN-SVLAN mappings.

Configuring an uplink policy

To configure an uplink policy to map each CVLAN to a unique SVLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a class and enter class view.	traffic classifier <i>tcl-name</i> [operator { and or }]	
3. Specify a CVLAN as the match criterion.	if-match customer-vlan-id <i>vlan-id</i>	Repeat these steps to configure one class for each CVLAN.
4. Return to system view.	quit	
5. Create a traffic behavior and enter traffic behavior view.	traffic behavior <i>behavior-name</i>	
6. Configure an SVLAN marking action.	remark service-vlan-id <i>vlan-id</i>	Repeat these steps to configure one behavior for each SVLAN.
7. Return to system view.	quit	
8. Create a QoS policy and enter QoS policy view.	qos policy <i>policy-name</i>	N/A
9. Associate the class with the behavior to map the CVLAN to the SVLAN.	classifier <i>tcl-name</i> behavior <i>behavior-name</i>	Repeat this step to create other CVLAN-to-SVLAN mappings.

Configuring a downlink policy

To configure a downlink policy to map SVLANs back to CVLANs:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a class and enter class view.	traffic classifier <i>tcl-name</i> [operator { and or }]	Repeat these steps to configure one class for each SVLAN.
3. Configure an SVLAN as the match criterion.	if-match service-vlan-id <i>vlan-id</i>	
4. Return to system view.	quit	
5. Create a traffic behavior and enter traffic behavior view.	traffic behavior <i>behavior-name</i>	Repeat these steps to configure a behavior for each CVLAN.
6. Configure a CVLAN marking action.	remark customer-vlan-id <i>vlan-id</i>	
7. Return to system view.	quit	
8. Create a QoS policy and enter QoS policy view.	qos policy <i>policy-name</i>	N/A
9. Associate the class with the behavior to map the SVLAN to the CVLAN.	classifier <i>tcl-name</i> behavior <i>behavior-name</i>	Repeat this step to create other CVLAN-to-SVLAN mappings.

Configuring the customer-side port

To configure the customer-side port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the port as a trunk port.	port link-type trunk	The default link type of an Ethernet port is access.
4. Assign the port to CVLANs and SVLANs.	port trunk permit vlan { <i>vlan-id-list</i> all }	By default, a trunk port is in only VLAN 1.
5. Enable basic QinQ.	qinq enable	By default, basic QinQ is disabled.
6. Apply the uplink policy to the incoming traffic.	qos apply policy <i>policy-name</i> inbound	N/A
7. Apply the downlink policy to the outgoing traffic.	qos apply policy <i>policy-name</i> outbound	N/A

Configuring the network-side port

To configure the network-side port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A

Step	Command	Remarks
3. Configure the port as a trunk port.	port link-type trunk	The default link type of ports is access.
4. Assign the port to SVLANs.	port trunk permit vlan { <i>vlan-id-list</i> all }	By default, a trunk port is in only VLAN 1.

Configuring many-to-one VLAN mapping

⚠ CAUTION:

Before changing VLAN mappings on a port, clear all DHCP snooping entries by using the **reset dhcp-snooping** command (see *Layer 3—IP Services Command Reference*).

Perform many-to-one VLAN mapping on campus switches (see [Figure 64](#)) to transmit the same type of traffic from different users in one VLAN.

Perform these tasks to configure many-to-one VLAN mapping:

Task	Description
Enabling DHCP snooping	Enables DHCP snooping globally (required).
Enabling ARP detection in SVLANs	Enables ARP detection in all SVLANs (required).
Configuring an uplink policy	Configures an uplink policy for the customer-side port (required).
Configuring the customer-side port	Configures VLAN and other settings required for many-to-one VLAN mapping (required).
Configuring the network-side port	Configures VLAN and other settings required for many-to-one VLAN mapping (required).

Configuration prerequisites

Before configuring many-to-one VLAN mapping:

- Make sure that all home users obtain IP addresses through DHCP. For how to assign IP addresses through DHCP, see *Layer 3—IP Services Configuration Guide*.
- Create CVLANs and SVLANs, and plan CVLANs-to-SVLAN mappings.

Enabling DHCP snooping

To enable DHCP snooping:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable DHCP snooping.	dhcp-snooping	Disabled by default.

Enabling ARP detection in SVLANs

The ARP detection function enables a switch to modify the VLAN attributes of ARP packets, which is impossible under the normal ARP packet processing procedure. For more information about ARP detection, see *Security Configuration Guide*.

To enable ARP detection in all SVLANs:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
3. Enable ARP detection.	arp detection enable	Disabled by default.

NOTE:

To defend against ARP attacks, enable ARP detection also in all CVLANs.

Configuring an uplink policy

To configure an uplink policy to map a group of CVLANs to one SVLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a class and enter class view.	traffic classifier <i>tcl-name</i> operator or	
3. Configure multiple CVLANs as match criteria.	if-match customer-vlan-id { <i>vlan-id-list</i> <i>vlan-id1</i> to <i>vlan-id2</i> }	Repeat these steps to configure one class for each group of CVLANs.
4. Return to system view.	quit	
5. Create a traffic behavior and enter traffic behavior view.	traffic behavior <i>behavior-name</i>	
6. Configure an SVLAN marking action.	remark service-vlan-id <i>vlan-id</i>	Repeat these steps to configure one behavior for each SVLAN.
7. Return to system view.	quit	
8. Create a QoS policy and enter QoS policy view.	qos policy <i>policy-name</i>	N/A
9. Map the CVLANs to the SVLAN by associating the class with the behavior.	classifier <i>tcl-name</i> behavior <i>behavior-name</i> mode dot1q-tag-manipulation	Repeat this step to create other CVLANs-to-SVLAN mappings.

Configuring the customer-side port

△ CAUTION:

Before applying a QoS policy to the customer-side port, enable customer-side QinQ on the port. Before disabling customer-side QinQ on the customer-side port, remove the QoS policy from the port first.

To configure the customer-side port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the port as a trunk port.	port link-type trunk	The default link type of an Ethernet port is access.
4. Assign the port to CVLANs and SVLANs.	port trunk permit vlan { <i>vlan-id-list</i> all }	By default, a trunk port is in only VLAN 1.
5. Enable customer-side QinQ.	qinq enable downlink	By default, customer-side QinQ is disabled on all ports.
6. Apply the uplink policy to the incoming traffic.	qos apply policy <i>policy-name</i> inbound	N/A

Configuring the network-side port

To configure the network-side port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the port as a trunk port.	port link-type trunk	The default link type of an Ethernet port is access.
4. Assign the port to SVLANs.	port trunk permit vlan { <i>vlan-id-list</i> all }	By default, a trunk port is in only VLAN 1.
5. Configure the port as a DHCP snooping trusted port.	dhcp-snooping trust	By default, all ports are DHCP snooping untrusted ports.
6. Configure the port as an ARP trusted port.	arp detection trust	By default, all ports are ARP untrusted ports.
7. Enable network-side QinQ.	qinq enable uplink	By default, network-side QinQ is disabled on all ports.

Configuring two-to-two VLAN mapping

Perform two-to-two VLAN mapping on an edge device that connects two SP networks, for example, on PE 3 in [Figure 65](#). Two-to-two VLAN mapping enables two remote sites in different VLANs to communicate at Layer 2 across two service provider networks that use different VLAN assignment schemes.

For the ease of description, the VLAN tags of the double-tagged frames that arrive at the customer-side port are called foreign CVLANs and SVLANs, and the VLAN tags marked by the edge device are called local CVLANs and SVLANs.

Perform these tasks to configure two-to-two VLAN mapping:

Task	Description
Configuring an uplink policy for the customer-side port	Replaces foreign SVLANs with local SVLANs for uplink traffic (required).
Configuring an uplink policy for the network-side port	Replaces foreign CVLANs with local CVLANs for uplink traffic (required).
Configuring a downlink policy for the customer-side port	Replaces local SVLANs and CVLANs with foreign SVLANs and CVLANs (required).
Configuring the customer-side port	Configures VLAN and other settings required for two-to-two VLAN mapping (required).
Configuring the network-side port	Configures VLAN and other settings required for two-to-two VLAN mapping (required).

Configuring an uplink policy for the customer-side port

The uplink policy on the customer-side port modifies the SVLAN ID of incoming traffic.

To configure an uplink policy for the customer-side port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a class and enter class view.	traffic classifier <i>tcl-name</i> [operator and]	
3. Specify a foreign CVLAN as a match criterion.	if-match customer-vlan-id <i>vlan-id</i>	Repeat these steps to create one class for each foreign CVLAN and SVLAN pair.
4. Specify a foreign SVLAN as a match criterion.	if-match service-vlan-id <i>vlan-id</i>	
5. Return to system view.	quit	
6. Create a traffic behavior and enter traffic behavior view.	traffic behavior <i>behavior-name</i>	
7. Configure an SVLAN marking action to replace the foreign SVLAN ID with a local SVLAN ID.	remark service-vlan-id <i>vlan-id</i>	Repeat these steps to configure one SVLAN marking action for each CVLAN and SVLAN pair.
8. Return to system view.	quit	
9. Create a QoS policy and enter QoS policy view.	qos policy <i>policy-name</i>	N/A
10. Associate the class with the behavior.	classifier <i>tcl-name</i> behavior <i>behavior-name</i>	Repeat this step to create other class-behavior associations.

Configuring an uplink policy for the network-side port

The uplink policy on the network-side port modifies the CVLAN ID of incoming traffic.

To configure an uplink policy for the network-side port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a class and enter class view.	traffic classifier <i>tcl-name</i> [operator and]	
3. Specify a foreign CVLAN as a match criterion.	if-match customer-vlan-id <i>vlan-id</i>	Repeat these steps to create one class for each local SVLAN and foreign CVLAN pair.
4. Specify a local SVLAN as a match criterion.	if-match service-vlan-id <i>vlan-id</i>	
5. Return to system view.	quit	
6. Create a traffic behavior and enter traffic behavior view.	traffic behavior <i>behavior-name</i>	Repeat these steps to configure one CVLAN marking action for each local SVLAN and foreign CVLAN pair.
7. Configure a CVLAN marking action to replace the foreign CVLAN ID with a local CVLAN ID.	remark customer-vlan-id <i>vlan-id</i>	
8. Return to system view.	quit	
9. Create a QoS policy and enter QoS policy view.	qos policy <i>policy-name</i>	N/A
10. Associate the class with the behavior.	classifier <i>tcl-name</i> behavior <i>behavior-name</i>	Repeat this step to create other class-behavior associations.

Configuring a downlink policy for the customer-side port

The downlink policy on the customer-side port replaces local SVLAN and CVLAN pairs with foreign SVLAN and CVLAN pairs.

To configure a downlink policy for the customer-side port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a class and enter class view.	traffic classifier <i>tcl-name</i> [operator and]	
3. Specify a local CVLAN as a match criterion.	if-match customer-vlan-id <i>vlan-id</i>	Repeat these steps to create one class for each local CVLAN and SVLAN pair.
4. Specify a local SVLAN as a match criterion.	if-match service-vlan-id <i>vlan-id</i>	
5. Return to system view.	quit	
6. Create a traffic behavior and enter traffic behavior view.	traffic behavior <i>behavior-name</i>	Repeat these steps to create one VLAN marking behavior for each local CVLAN and SVLAN pair.
7. Configure a CVLAN marking action to replace the local CVLAN ID with a foreign CVLAN ID.	remark customer-vlan-id <i>vlan-id</i>	
8. Configure an SVLAN marking action to replace the local SVLAN ID with a foreign SVLAN ID.	remark service-vlan-id <i>vlan-id</i>	
9. Return to system view.	quit	

Step	Command	Remarks
10. Create a QoS policy and enter QoS policy view.	qos policy <i>policy-name</i>	N/A
11. Associate the class with the behavior.	classifier <i>tcl-name</i> behavior <i>behavior-name</i>	Repeat this step to create other class-behavior associations.

Configuring the customer-side port

To configure the customer-side port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the port as a trunk port.	port link-type trunk	The default link type of an Ethernet port is access.
4. Assign the port to the local SVLANs.	port trunk permit vlan { <i>vlan-id-list</i> all }	By default, a trunk port is in only VLAN 1.
5. Apply the uplink policy configured for the customer-side port to the incoming traffic.	qos apply policy <i>policy-name</i> inbound	N/A
6. Apply the downlink policy configured for the customer-side port to the outgoing traffic.	qos apply policy <i>policy-name</i> outbound	N/A

Configuring the network-side port

To configure the network-side port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the port as a trunk port.	port link-type trunk	The default link type of an Ethernet port is access.
4. Assign the port to local SVLANs.	port trunk permit vlan { <i>vlan-id-list</i> all }	By default, a trunk port is in only VLAN 1.
5. Apply the uplink policy for the network-side port to the outgoing traffic.	qos apply policy <i>policy-name</i> outbound	N/A

VLAN mapping configuration examples

One-to-one and many-to-one VLAN mapping configuration example

Network requirements

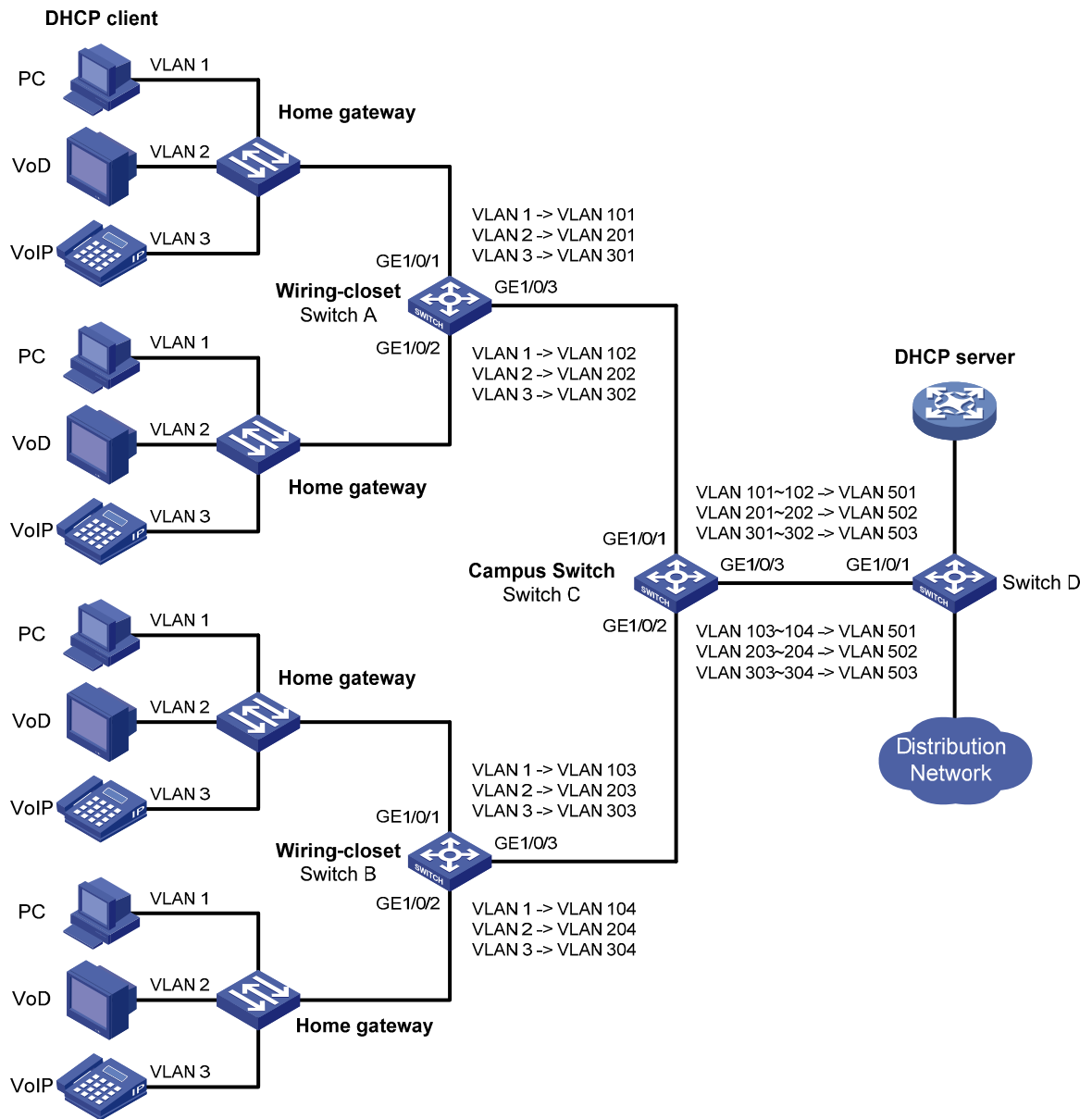
As shown in [Figure 70](#):

- Each home is offered PC, VoD, and VoIP services, connects to a wiring-closet switch through the home gateway, and obtains the IP address through DHCP.
- VLAN 1 is assigned for PC traffic, VLAN 2 is assigned for VoD traffic, and VLAN 3 is assigned for VoIP service traffic on the home gateways.

Assign one VLAN for each type of traffic from each user on the wiring-closet switches, so that traffic of the same service type from different homes is isolated.

Perform many-to-one VLAN mapping on the campus switch (Switch C) to transmit the same type of traffic from different homes in one VLAN: use VLAN 501 for PC traffic, VLAN 502 for VoD traffic, and VLAN 503 for VoIP traffic, so that the VLAN resources are saved.

Figure 70 Network diagram



Configuration procedure

1. Configure Switch A:

Create the CVLANs and the SVLANs.

```
<SwitchA> system-view
[SwitchA] vlan 2 to 3
[SwitchA] vlan 101 to 102
[SwitchA] vlan 201 to 202
[SwitchA] vlan 301 to 302
```

Configure uplink policies **p1** and **p2** to enable one SVLAN to transmit one service for one customer.

```
[SwitchA] traffic classifier c1
[SwitchA-classifier-c1] if-match customer-vlan-id 1
```



```

[SwitchA-classifier-c1] traffic classifier c2
[SwitchA-classifier-c2] if-match customer-vlan-id 2
[SwitchA-classifier-c2] traffic classifier c3
[SwitchA-classifier-c3] if-match customer-vlan-id 3
[SwitchA-classifier-c3] quit
[SwitchA] traffic behavior b1
[SwitchA-behavior-b1] remark service-vlan-id 101
[SwitchA-behavior-b1] traffic behavior b2
[SwitchA-behavior-b2] remark service-vlan-id 201
[SwitchA-behavior-b2] traffic behavior b3
[SwitchA-behavior-b3] remark service-vlan-id 301
[SwitchA-behavior-b3] traffic behavior b4
[SwitchA-behavior-b4] remark service-vlan-id 102
[SwitchA-behavior-b4] traffic behavior b5
[SwitchA-behavior-b5] remark service-vlan-id 202
[SwitchA-behavior-b5] traffic behavior b6
[SwitchA-behavior-b6] remark service-vlan-id 302
[SwitchA-behavior-b6] quit
[SwitchA] qos policy p1
[SwitchA-policy-p1] classifier c1 behavior b1
[SwitchA-policy-p1] classifier c2 behavior b2
[SwitchA-policy-p1] classifier c3 behavior b3
[SwitchA-policy-p1] quit
[SwitchA] qos policy p2
[SwitchA-policy-p2] classifier c1 behavior b4
[SwitchA-policy-p2] classifier c2 behavior b5
[SwitchA-policy-p2] classifier c3 behavior b6
[SwitchA-policy-p2] quit
# Configure downlink policies p11 and p22 to map the SVLANs back to the CVLANs.
[SwitchA] traffic classifier c11
[SwitchA-classifier-c11] if-match service-vlan-id 101
[SwitchA-classifier-c11] traffic classifier c22
[SwitchA-classifier-c22] if-match service-vlan-id 201
[SwitchA-classifier-c22] traffic classifier c33
[SwitchA-classifier-c33] if-match service-vlan-id 301
[SwitchA-classifier-c33] traffic classifier c44
[SwitchA-classifier-c44] if-match service-vlan-id 102
[SwitchA-classifier-c44] traffic classifier c55
[SwitchA-classifier-c55] if-match service-vlan-id 202
[SwitchA-classifier-c55] traffic classifier c66
[SwitchA-classifier-c66] if-match service-vlan-id 302
[SwitchA-classifier-c66] quit
[SwitchA] traffic behavior b11
[SwitchA-behavior-b11] remark customer-vlan-id 1
[SwitchA-behavior-b11] traffic behavior b22
[SwitchA-behavior-b22] remark customer-vlan-id 2
[SwitchA-behavior-b22] traffic behavior b33
[SwitchA-behavior-b33] remark customer-vlan-id 3

```

```

[SwitchA-behavior-b33] quit
[SwitchA] qos policy p11
[SwitchA-policy-p11] classifier c11 behavior b11
[SwitchA-policy-p11] classifier c22 behavior b22
[SwitchA-policy-p11] classifier c33 behavior b33
[SwitchA-policy-p11] quit
[SwitchA] qos policy p22
[SwitchA-policy-p22] classifier c44 behavior b11
[SwitchA-policy-p22] classifier c55 behavior b22
[SwitchA-policy-p22] classifier c66 behavior b33
[SwitchA-policy-p22] quit
# Assign customer-side port GigabitEthernet 1/0/1 to CVLANs 1 to 3, and SVLANs 101, 201,
and 301, and enable basic QinQ, and apply uplink policy p1 to the incoming traffic and
downlink policy p11 to the outgoing traffic.
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] port link-type trunk
[SwitchA-GigabitEthernet1/0/1] port trunk permit vlan 1 2 3 101 201 301
[SwitchA-GigabitEthernet1/0/1] qinq enable
[SwitchA-GigabitEthernet1/0/1] qos apply policy p1 inbound
[SwitchA-GigabitEthernet1/0/1] qos apply policy p11 outbound
[SwitchA-GigabitEthernet1/0/1] quit
# Assign customer-side port GigabitEthernet 1/0/2 to CVLANs 1 to 3, and SVLANs 102, 202,
and 302, enable basic QinQ, and apply uplink policy p2 to the incoming traffic and downlink
policy p22 to the outgoing traffic.
[SwitchA] interface gigabitethernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] port link-type trunk
[SwitchA-GigabitEthernet1/0/2] port trunk permit vlan 1 2 3 102 202 302
[SwitchA-GigabitEthernet1/0/2] qinq enable
[SwitchA-GigabitEthernet1/0/2] qos apply policy p2 inbound
[SwitchA-GigabitEthernet1/0/2] qos apply policy p22 outbound
[SwitchA-GigabitEthernet1/0/2] quit
# Assign network-side port GigabitEthernet 1/0/3 to all SVLANs.
[SwitchA] interface gigabitethernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] port link-type trunk
[SwitchA-GigabitEthernet1/0/3] port trunk permit vlan 101 201 301 102 202 302

```

2. Configure Switch B:

Configure Switch B by using the same procedure as on Switch A.

3. Configure Switch C:

Enable DHCP snooping.

```

<SwitchC> system-view
[SwitchC] dhcp-snooping

```

Create the CVLANs and SVLANs, and enable ARP detection in each VLAN.

```

[SwitchC] vlan 101
[SwitchC-vlan101] arp detection enable
[SwitchC-vlan101] vlan 201
[SwitchC-vlan201] arp detection enable
[SwitchC-vlan201] vlan 301

```

```

[SwitchC-vlan301] arp detection enable
[SwitchC-vlan301] vlan 102
[SwitchC-vlan102] arp detection enable
[SwitchC-vlan102] vlan 202
[SwitchC-vlan202] arp detection enable
[SwitchC-vlan202] vlan 302
[SwitchC-vlan302] arp detection enable
[SwitchC-vlan302] vlan 103
[SwitchC-vlan103] arp detection enable
[SwitchC-vlan103] vlan 203
[SwitchC-vlan203] arp detection enable
[SwitchC-vlan203] vlan 303
[SwitchC-vlan303] arp detection enable
[SwitchC-vlan303] vlan 104
[SwitchC-vlan104] arp detection enable
[SwitchC-vlan104] vlan 204
[SwitchC-vlan204] arp detection enable
[SwitchC-vlan204] vlan 304
[SwitchC-vlan304] arp detection enable
[SwitchC-vlan304] vlan 501
[SwitchC-vlan501] arp detection enable
[SwitchC-vlan501] vlan 502
[SwitchC-vlan502] arp detection enable
[SwitchC-vlan502] vlan 503
[SwitchC-vlan503] arp detection enable
[SwitchC-vlan503] quit

```

Configure uplink policies p1 and p2 to enable one SVLAN to transmit the same type of traffic from different customers.

```

[SwitchC] traffic classifier c1
[SwitchC-classifier-c1] if-match customer-vlan-id 101 to 102
[SwitchC-classifier-c1] traffic classifier c2
[SwitchC-classifier-c2] if-match customer-vlan-id 201 to 202
[SwitchC-classifier-c2] traffic classifier c3
[SwitchC-classifier-c3] if-match customer-vlan-id 301 to 302
[SwitchC-classifier-c3] traffic classifier c4
[SwitchC-classifier-c4] if-match customer-vlan-id 103 to 104
[SwitchC-classifier-c4] traffic classifier c5
[SwitchC-classifier-c5] if-match customer-vlan-id 203 to 204
[SwitchC-classifier-c5] traffic classifier c6
[SwitchC-classifier-c6] if-match customer-vlan-id 303 to 304
[SwitchC-classifier-c6] quit
[SwitchC] traffic behavior b1
[SwitchC-behavior-b1] remark service-vlan-id 501
[SwitchC-behavior-b1] traffic behavior b2
[SwitchC-behavior-b2] remark service-vlan-id 502
[SwitchC-behavior-b2] traffic behavior b3
[SwitchC-behavior-b3] remark service-vlan-id 503
[SwitchC-behavior-b3] quit

```

```

[SwitchC] qos policy p1
[SwitchC-policy-p1] classifier c1 behavior b1 mode dot1q-tag-manipulation
[SwitchC-policy-p1] classifier c2 behavior b2 mode dot1q-tag-manipulation
[SwitchC-policy-p1] classifier c3 behavior b3 mode dot1q-tag-manipulation
[SwitchC-policy-p1] quit
[SwitchC] qos policy p2
[SwitchC-policy-p2] classifier c4 behavior b1 mode dot1q-tag-manipulation
[SwitchC-policy-p2] classifier c5 behavior b2 mode dot1q-tag-manipulation
[SwitchC-policy-p2] classifier c6 behavior b3 mode dot1q-tag-manipulation
[SwitchC-policy-p2] quit
# Assign customer-side port GigabitEthernet 1/0/1 to CVLANs 101, 201, 301, 102, 202, 302,
and SVLANs 501 to 503. On this port, also enable customer-side QinQ, and apply uplink policy
p1 to the incoming traffic.
[SwitchC] interface gigabitethernet 1/0/1
[SwitchC-GigabitEthernet1/0/1] port link-type trunk
[SwitchC-GigabitEthernet1/0/1] port trunk permit vlan 101 201 301 102 202 302 501 502
503
[SwitchC-GigabitEthernet1/0/1] qinq enable downlink
[SwitchC-GigabitEthernet1/0/1] qos apply policy p1 inbound
[SwitchC-GigabitEthernet1/0/1] quit
# Assign customer-side port GigabitEthernet 1/0/2 to CVLANs 103, 203, 303, 104, 204, 304,
and SVLANs 501 to 503. On this port, also enable customer-side QinQ, and apply uplink policy
p2 to the incoming traffic.
[SwitchC] interface gigabitethernet 1/0/2
[SwitchC-GigabitEthernet1/0/2] port link-type trunk
[SwitchC-GigabitEthernet1/0/2] port trunk permit vlan 103 203 303 104 204 304 501 502
503
[SwitchC-GigabitEthernet1/0/2] qinq enable downlink
[SwitchC-GigabitEthernet1/0/2] qos apply policy p2 inbound
[SwitchC-GigabitEthernet1/0/2] quit
# Assign network-side GigabitEthernet 1/0/3 to SVLANs 501 to 503, set the port as a DHCP and
ARP trusted port, and enable network-side QinQ.
[SwitchC] interface gigabitethernet 1/0/3
[SwitchC-GigabitEthernet1/0/3] port link-type trunk
[SwitchC-GigabitEthernet1/0/3] port trunk permit vlan 501 502 503
[SwitchC-GigabitEthernet1/0/3] dhcp-snooping trust
[SwitchC-GigabitEthernet1/0/3] arp detection trust
[SwitchC-GigabitEthernet1/0/3] qinq enable uplink

```

4. Configure Switch D:

```

# Enable DHCP snooping.
<SwitchD> system-view
[SwitchD] dhcp-snooping
# Assign port GigabitEthernet 1/0/1 to SVLANs 501 to 503.
[SwitchD] interface gigabitethernet 1/0/1
[SwitchD-GigabitEthernet1/0/1] port link-type trunk
[SwitchD-GigabitEthernet1/0/1] port trunk permit vlan 501 502 503

```

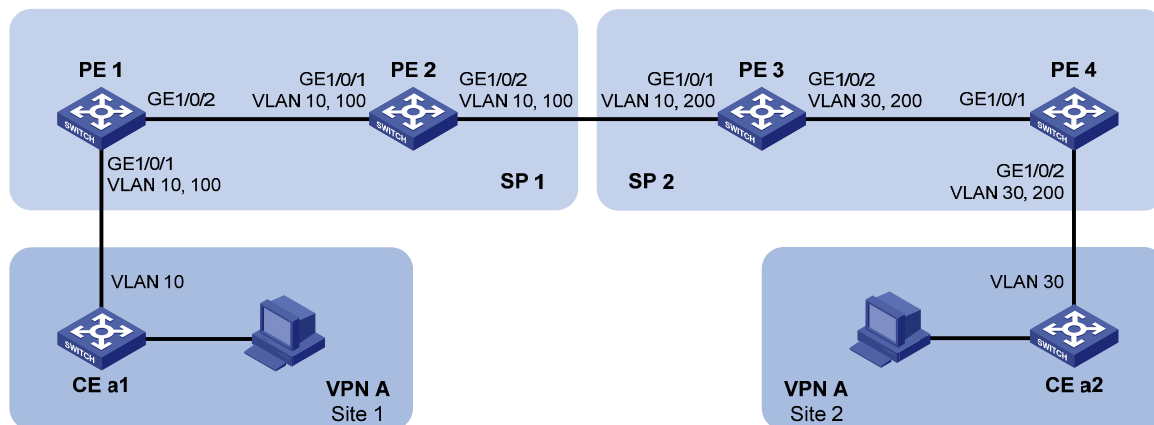
Two-to-two VLAN mapping configuration example

Network requirements

As shown in Figure 71, two VPN A branches, Site 1 and Site 2, are in VLAN 10 and VLAN 30, respectively. The two sites use different VPN access services of different service providers, SP 1 and SP 2. SP 1 assigns VLAN 100 for Site 1, and SP 2 assigns VLAN 200 for Site 2.

Configure two-to-two VLAN mappings to enable the two branches to communicate across networks SP 1 and SP 2.

Figure 71 Network diagram



Configuration procedure

1. Configure PE 1:

Configure QinQ function on GigabitEthernet 1/0/1 to add outer VLAN tag 100 to the traffic tagged with VLAN 10.

```
<PE1> system-view
[PE1] interface gigabitEthernet 1/0/1
[PE1-GigabitEthernet1/0/1] port access vlan 100
[PE1-GigabitEthernet1/0/1] qinq enable
[PE1-GigabitEthernet1/0/1] quit
# Configure the uplink port GigabitEthernet 1/0/2 to permit frames of VLAN 100 to pass through.
[PE1] interface gigabitEthernet 1/0/2
[PE1-GigabitEthernet1/0/2] port link-type trunk
[PE1-GigabitEthernet1/0/2] port trunk permit vlan 100
```

2. Configure PE 2:

Set port GigabitEthernet 1/0/1 as a trunk port, and assign it to VLAN 100.

```
<PE2> system-view
[PE2] interface gigabitEthernet 1/0/1
[PE2-GigabitEthernet1/0/1] port link-type trunk
[PE2-GigabitEthernet1/0/1] port trunk permit vlan 100
[PE2-GigabitEthernet1/0/1] quit
# Set port GigabitEthernet 1/0/2 as a trunk port, and assign it to VLAN 100.
[PE2] interface gigabitEthernet 1/0/2
[PE2-GigabitEthernet1/0/2] port link-type trunk
[PE2-GigabitEthernet1/0/2] port trunk permit vlan 100
```

3. Configure PE 3:

Configure an uplink policy **down_uplink** for customer-side port GigabitEthernet 1/0/1 to substitute SVLAN ID 200 for the SVLAN ID in the incoming traffic tagged with CVLAN 10 and SVLAN 100.

```
<PE3> system-view
[PE3] traffic classifier down_uplink
[PE3-classifier-down_uplink] if-match customer-vlan-id 10
[PE3-classifier-down_uplink] if-match service-vlan-id 100
[PE3-classifier-down_uplink] quit
[PE3] traffic behavior down_uplink
[PE3-behavior-down_uplink] remark service-vlan-id 200
[PE3-behavior-down_uplink] quit
[PE3] qos policy down_uplink
[PE3-qospolicy-down_uplink] classifier down_uplink behavior down_uplink
[PE3-qospolicy-down_uplink] quit
```

Configure a downlink policy **down_downlink** for customer-side port GigabitEthernet 1/0/1 to substitute CVLAN 10 and SVLAN 100 for traffic tagged with CVLAN 30 and SVLAN 200.

```
[PE3] traffic classifier down_downlink
[PE3-classifier-down_downlink] if-match customer-vlan-id 30
[PE3-classifier-down_downlink] if-match service-vlan-id 200
[PE3-classifier-down_downlink] quit
[PE3] traffic behavior down_downlink
[PE3-behavior-down_downlink] remark customer-vlan-id 10
[PE3-behavior-down_downlink] remark service-vlan-id 100
[PE3-behavior-down_downlink] quit
[PE3] qos policy down_downlink
[PE3-qospolicy-down_downlink] classifier down_downlink behavior down_downlink
[PE3-qospolicy-down_downlink] quit
```

Configure an uplink policy **up_uplink** for network-side port GigabitEthernet 1/0/2 to substitute CVLAN 30 for the CVLAN ID of the outgoing traffic tagged with CVLAN 10 and SVLAN 200.

```
[PE3] traffic classifier up_uplink
[PE3-classifier-up_uplink] if-match customer-vlan-id 10
[PE3-classifier-up_uplink] if-match service-vlan-id 200
[PE3-classifier-up_uplink] quit
[PE3] traffic behavior up_uplink
[PE3-behavior-up_uplink] remark customer-vlan-id 30
[PE3-behavior-up_uplink] quit
[PE3] qos policy up_uplink
[PE3-qospolicy-up_uplink] classifier up_uplink behavior up_uplink
[PE3-qospolicy-up_uplink] quit
```

Set customer-side port GigabitEthernet 1/0/1 as a trunk port, assign it to VLAN 200, and apply uplink policy **down_uplink** to the incoming traffic and downlink policy **down_downlink** to the outgoing traffic on the port.

```
[PE3] interface gigabitethernet 1/0/1
[PE3-GigabitEthernet1/0/1] port link-type trunk
[PE3-GigabitEthernet1/0/1] port trunk permit vlan 200
[PE3-GigabitEthernet1/0/1] qos apply policy down_uplink inbound
[PE3-GigabitEthernet1/0/1] qos apply policy down_downlink outbound
```

```
[PE3-GigabitEthernet1/0/1] quit
```

Set network-side port GigabitEthernet 1/0/2 as a trunk port, assign it to VLAN 200, and apply uplink policy **up_uplink** to the outgoing traffic on the port.

```
[PE3] interface gigabitethernet 1/0/2
```

```
[PE3-GigabitEthernet1/0/2] port link-type trunk
```

```
[PE3-GigabitEthernet1/0/2] port trunk permit vlan 200
```

```
[PE3-GigabitEthernet1/0/2] qos apply policy up_uplink outbound
```

```
[PE3-GigabitEthernet1/0/2] quit
```

4. Configure PE 4:

Configure QinQ function on GigabitEthernet 1/0/2 to add outer VLAN tag 200 to the traffic tagged with VLAN 30.

```
<DeviceD> system-view
```

```
[DeviceD] interface gigabitethernet 1/0/2
```

```
[DeviceD-GigabitEthernet1/0/2] port access vlan 200
```

```
[DeviceD-GigabitEthernet1/0/2] qinq enable
```

Configure GigabitEthernet 1/0/1 to permit frames of VLAN 200 to pass through.

```
[DeviceD] interface gigabitethernet 1/0/1
```

```
[DeviceD-GigabitEthernet1/0/1] port link-type trunk
```

```
[DeviceD-GigabitEthernet1/0/1] port trunk permit vlan 200
```

Configuring LLDP

Overview

Background

In a heterogeneous network, a standard configuration exchange platform ensures that different types of network devices from different vendors can discover one another and exchange configuration for the sake of interoperability and management.

The IETF drafted the Link Layer Discovery Protocol (LLDP) in IEEE 802.1AB. The protocol operates on the data link layer to exchange device information between directly connected devices. With LLDP, a device sends local device information (including its major functions, management IP address, device ID, and port ID) as TLV (type, length, and value) triplets in LLDP Data Units (LLDPDUs) to the directly connected devices. At the same time, the device stores the device information received in LLDPDUs sent from the LLDP neighbors in a standard management information base (MIB). For more information about MIBs, see *Network Management and Monitoring Configuration Guide*. LLDP enables a network management system to quickly detect and identify Layer 2 network topology changes.

Basic concepts

LLDPDU formats

LLDP sends device information in LLDPDUs. LLDPDUs are encapsulated in Ethernet II or Subnetwork Access Protocol (SNAP) frames.

1. Ethernet II-encapsulated LLDPDU format

Figure 72 Ethernet II-encapsulated LLDPDU format

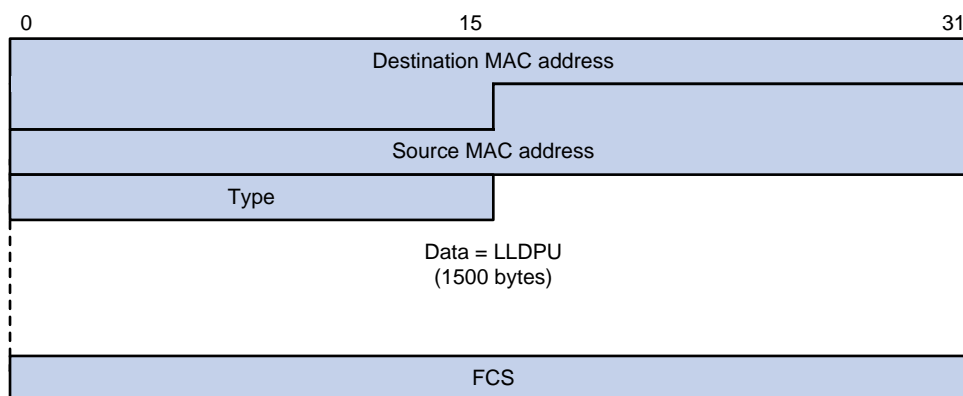


Table 22 Fields in an Ethernet II-encapsulated LLDPDU

Field	Description
Destination MAC address	MAC address to which the LLDPDU is advertised. It is fixed to 0x0180-C200-000E, a multicast MAC address.

Field	Description
Source MAC address	MAC address of the sending port. If the port does not have a MAC address, the MAC address of the sending bridge is used.
Type	Ethernet type for the upper layer protocol. It is 0x88CC for LLDP.
Data	LLDPDU.
FCS	Frame check sequence, a 32-bit CRC value used to determine the validity of the received Ethernet frame.

2. SNAP-encapsulated LLDPDU format

Figure 73 SNAP-encapsulated LLDPDU format

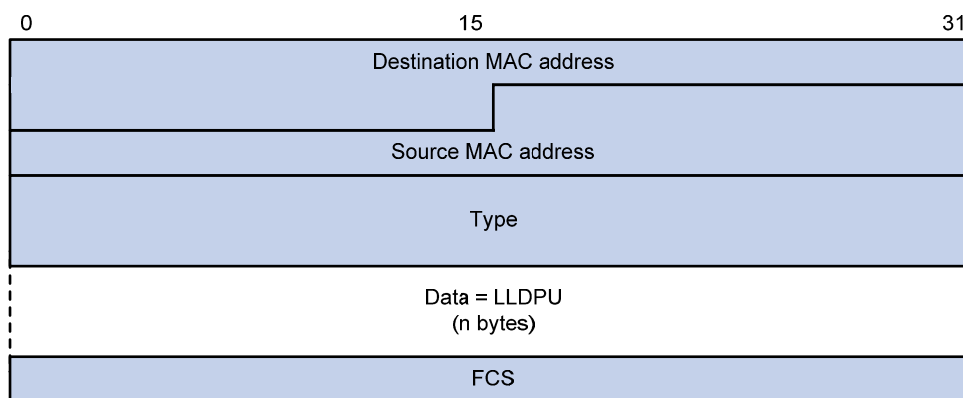


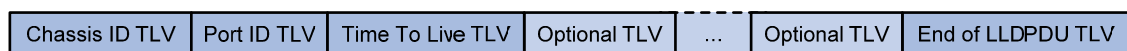
Table 23 Fields in a SNAP-encapsulated LLDPDU

Field	Description
Destination MAC address	MAC address to which the LLDPDU is advertised. It is fixed at 0x0180-C200-000E, a multicast MAC address.
Source MAC address	MAC address of the sending port.
Type	SNAP type for the upper layer protocol. It is 0xAAAA-0300-0000-88CC for LLDP.
Data	LLDPDU.
FCS	Frame check sequence, a 32-bit CRC value used to determine the validity of the received Ethernet frame.

LLDPDU_s

LLDP uses LLDPDU_s to exchange information. An LLDPDU comprises multiple TLV sequences. Each TLV carries a type of device information, as shown in [Figure 74](#).

Figure 74 LLDPDU encapsulation format



An LLDPDU can carry up to 28 types of TLVs. Mandatory TLVs include Chassis ID TLV, Port ID TLV, Time To Live TLV, and End of LLDPDU TLV. Other TLVs are optional.

TLVs

TLVs are type, length, and value sequences that carry information elements. The type field identifies the type of information, the length field measures the length of the information field in octets, and the value field contains the information itself.

LLDPDU TLVs fall into the following categories:

- Basic management TLVs
- Organizationally (IEEE 802.1 and IEEE 802.3) specific TLVs
- LLDP-MED (media endpoint discovery) TLVs

Basic management TLVs are essential to device management. Organizationally specific TLVs and LLDP-MED TLVs are used for enhanced device management; they are defined by standardization or other organizations and are optional to LLDPDUs.

1. Basic management TLVs

Table 24 lists the basic management TLV types. Some of them are mandatory to LLDPDUs, that is, must be included in every LLDPDU.

Table 24 Basic management TLVs

Type	Description	Remarks
Chassis ID	Specifies the bridge MAC address of the sending device	
Port ID	Specifies the ID of the sending port If the LLDPDU carries LLDP-MED TLVs, the port ID TLV carries the MAC address of the sending port or the bridge MAC if the port does not have a MAC address. If the LLDPDU carries no LLDP-MED TLVs, the port ID TLV carries the port name.	Mandatory
Time To Live	Specifies the life of the transmitted information on the receiving device	
End of LLDPDU	Marks the end of the TLV sequence in the LLDPDU	
Port Description	Specifies the port description of the sending port	
System Name	Specifies the assigned name of the sending device	
System Description	Specifies the description of the sending device	
System Capabilities	Identifies the primary functions of the sending device and the enabled primary functions	Optional
Management Address	Specifies the management address, and the interface number and object identifier (OID) associated with the address	

2. IEEE 802.1 organizationally specific TLVs

Table 25 IEEE 802.1 organizationally specific TLVs

Type	Description
Port VLAN ID	Specifies the port's VLAN identifier (PVID). An LLDPDU carries only one TLV of this type.
Port And Protocol VLAN ID	Indicates whether the device supports protocol VLANs and, if so, what VLAN IDs these protocols will be associated with. An LLDPDU can carry multiple different TLVs of this type.

Type	Description
VLAN Name	Specifies the textual name of any VLAN to which the port belongs. An LLDPDU can carry multiple different TLVs of this type.
Protocol Identity	Indicates protocols supported on the port. An LLDPDU can carry multiple different TLVs of this type.

NOTE:

- HP devices support only receiving protocol identity TLVs.
- Layer 3 Ethernet ports do not support IEEE 802.1 organizationally specific TLVs.

3. IEEE 802.3 organizationally specific TLVs

Table 26 IEEE 802.3 organizationally specific TLVs

Type	Description
MAC/PHY Configuration/Status	Contains the bit-rate and duplex capabilities of the sending port, support for auto negotiation, enabling status of auto negotiation, and the current rate and duplex mode.
Power Via MDI	Contains the power supply capability of the port, including the Power over Ethernet (PoE) type, which can be Power Sourcing Equipment (PSE) or Powered Device (PD), PoE mode, whether PSE power supply is supported, whether PSE power supply is enabled, and whether the PoE mode is controllable.
Link Aggregation	Indicates the aggregation capability of the port (whether the link is capable of being aggregated), and the aggregation status (whether the link is in an aggregation).
Maximum Frame Size	Indicates the supported maximum frame size. It is now the maximum transmission unit (MTU) of the port.
Power Stateful Control	Indicates the power state control configured on the sending port, including the power type of the PSE or PD, PoE sourcing and receiving priority, and PoE sourcing and receiving power.

NOTE:

The Power Stateful Control TLV is defined in IEEE P802.3at D1.0. The later versions no longer support this TLV. HP devices send this type of TLVs only after receiving them.

LLDP-MED TLVs

LLDP-MED TLVs provide multiple advanced applications for voice over IP (VoIP), such as basic configuration, network policy configuration, and address and directory management. LLDP-MED TLVs provide a cost-effective and easy-to-use solution for deploying voice devices in Ethernet. LLDP-MED TLVs are shown in [Table 27](#).

Table 27 LLDP-MED TLVs

Type	Description
LLDP-MED Capabilities	Allows a network device to advertise the LLDP-MED TLVs that it supports.

Type	Description
Network Policy	Allows a network device or terminal device to advertise the VLAN ID of the specific port, the VLAN type, and the Layer 2 and Layer 3 priorities for specific applications.
Extended Power-via-MDI	Allows a network device or terminal device to advertise power supply capability. This TLV is an extension of the Power Via MDI TLV.
Hardware Revision	Allows a terminal device to advertise its hardware version.
Firmware Revision	Allows a terminal device to advertise its firmware version.
Software Revision	Allows a terminal device to advertise its software version.
Serial Number	Allows a terminal device to advertise its serial number.
Manufacturer Name	Allows a terminal device to advertise its vendor name.
Model Name	Allows a terminal device to advertise its model name.
Asset ID	Allows a terminal device to advertise its asset ID. The typical case is that the user specifies the asset ID for the endpoint to facilitate directory management and asset tracking.
Location Identification	Allows a network device to advertise the appropriate location identifier information for a terminal device to use in the context of location-based applications.

Management address

The network management system uses the management address of a device to identify and manage the device for topology maintenance and network management. The management address TLV encapsulates the management address.

How LLDP works

Operating modes of LLDP

LLDP can operate in one of the following modes:

- **TxRx mode**—A port in this mode sends and receives LLDPDUs.
- **Tx mode**—A port in this mode only sends LLDPDUs.
- **Rx mode**—A port in this mode only receives LLDPDUs.
- **Disable mode**—A port in this mode does not send or receive LLDPDUs.

Each time the LLDP operating mode of a port changes, its LLDP protocol state machine re-initializes. A re-initialization delay, which is user configurable, prevents LLDP from being initialized too frequently at times of frequent changes to the operating mode. With this delay configured, before a port can initialize LLDP, it must wait for the specified interval after the LLDP operating mode changes.

Transmitting LLDPDUs

An LLDP-enabled port operating in TxRx mode or Tx mode sends LLDPDUs to its directly connected devices both periodically and when the local configuration changes. To prevent LLDPDUs from overwhelming the network during times of frequent changes to local device information, an interval is introduced between two successive LLDPDUs.

This interval is shortened to 1 second in either of the following cases:

- A new neighbor is discovered. A new LLDPDU is received and carries device information new to the local device.
- The LLDP operating mode of the port changes from Disable or Rx to TxRx or Tx.

This is the fast sending mechanism of LLDP. With this mechanism, a specific number of LLDPDUs are sent successively at 1-second intervals, to help LLDP neighbors discover the local device as soon as possible. Then, the normal LLDPDU transmit interval resumes.

Receiving LLDPDUs

An LLDP-enabled port that is operating in TxRx mode or Rx mode checks the validity of TLVs carried in every received LLDPDU. If valid, the information is saved and an aging timer is set for it based on the time to live (TTL) value in the Time to Live TLV carried in the LLDPDU. If the TTL value is zero, the information ages out immediately.

Protocols and standards

- IEEE 802.1AB-2005, *Station and Media Access Control Connectivity Discovery*
- ANSI/TIA-1057, *Link Layer Discovery Protocol for Media Endpoint Devices*

LLDP configuration task list

When you configure LLDP, follow these guidelines:

- LLDP-related configurations made in Layer 2/Layer 3 Ethernet interface view take effect only on the current port, and those made in port group view take effect on all ports in the current port group.
- Layer 3 interfaces refer to route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see "[Configuring Ethernet interfaces](#)").
- The 5500 SI Switches Series does not support Layer 3 Ethernet interfaces.

Complete these tasks to configure LLDP:

Task	Remarks	
Enabling LLDP	Required	
Setting the LLDP operating mode	Optional	
Setting the LLDP re-initialization delay	Optional	
Performing basic LLDP configuration	Enabling LLDP polling	Optional
	Configuring the advertisable TLVs	Optional
	Configuring the management address and its encoding format	Optional
	Setting other LLDP parameters	Optional
	Setting an encapsulation format for LLDPDUs	Optional
Configuring CDP compatibility	Optional	
Enabling LLDP to automatically discover IP phones	Optional	
Configuring LLDP to advertise a specific voice VLAN	Optional	
Dynamically advertising server-assigned VLANs through LLDP	Optional	
Configuring LLDP trapping	Optional	

Performing basic LLDP configuration

Enabling LLDP

To make LLDP take effect on specific ports, you must enable LLDP both globally and on these ports.

To enable LLDP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable LLDP globally.	lldp enable	By default, LLDP is globally enabled.
3. Enter Ethernet interface view or port group view.	<ul style="list-style-type: none">Enter Layer 2/Layer 3 Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i>Enter port group view: port-group manual <i>port-group-name</i>	Use either command.
4. Enable LLDP.	lldp enable	Optional. By default, LLDP is enabled on a port.

Setting the LLDP operating mode

LLDP can operate in one of the following modes.

- **TxRx mode**—A port in this mode sends and receives LLDPDUs.
- **Tx mode**—A port in this mode only sends LLDPDUs.
- **Rx mode**—A port in this mode only receives LLDPDUs.
- **Disable mode**—A port in this mode does not send or receive LLDPDUs.

To set the LLDP operating mode:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view or port group view.	<ul style="list-style-type: none">Enter Layer 2/Layer 3 Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i>Enter port group view: port-group manual <i>port-group-name</i>	Use either command.
3. Set the LLDP operating mode.	lldp admin-status { disable rx tx txrx }	Optional. TxRx by default.

Setting the LLDP re-initialization delay

When LLDP operating mode changes on a port, the port initializes the protocol state machines after a certain delay. By adjusting the LLDP re-initialization delay, you can avoid frequent initializations caused by frequent changes to the LLDP operating mode on a port.

To set the LLDP re-initialization delay for ports:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the LLDP re-initialization delay.	lldp timer reinit-delay <i>delay</i>	Optional. 2 seconds by default.

Enabling LLDP polling

With LLDP polling enabled, a device periodically searches for local configuration changes. On detecting a configuration change, the device sends LLDPDUs to inform neighboring devices of the change.

To enable LLDP polling:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2/Layer 3 Ethernet interface view: interface <i>interface-type interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Enable LLDP polling and set the polling interval.	lldp check-change-interval <i>interval</i>	Disabled by default.

Configuring the advertisable TLVs

To configure the advertisable LLDPDU TLVs on the specified port or ports:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2/Layer 3 Ethernet interface view: interface <i>interface-type interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Configure the advertisable TLVs (Layer 2 Ethernet interface view or port group view).	lldp tlv-enable { basic-tlv { all port-description system-capability system-description system-name } dot1-tlv { all port-vlan-id protocol-vlan-id [<i>vlan-id</i>] vlan-name [<i>vlan-id</i>] } dot3-tlv { all link-aggregation mac-physic max-frame-size power } med-tlv { all capability inventory location-id { civic-address <i>device-type country-code</i> { <i>ca-type ca-value</i> } &<1-10> elin-address <i>tel-number</i> } network-policy power-over-ethernet } }	Optional. By default, all types of LLDP TLVs except location identification TLVs are advertisable on a Layer 2 Ethernet port.

Step	Command	Remarks
4. Configure the advertisable TLVs (Layer 3 Ethernet interface view).	lldp tlv-enable { basic-tlv { all port-description system-capability system-description system-name } dot3-tlv { all link-aggregation mac-physic max-frame-size power } med-tlv { all capability inventory location-id { civic-address <i>device-type country-code</i> { <i>ca-type ca-value</i> }&<1-10> elin-address <i>tel-number</i> } power-over-ethernet } }	Optional. By default, all types of LLDP TLVs, except IEEE 802.1 organizationally specific TLVs, network policy TLVs, and location identification TLVs, are advertisable on a Layer 3 Ethernet port.

Configuring the management address and its encoding format

LLDP encodes management addresses in numeric or character string format in management address TLVs.

By default, management addresses are encoded in numeric format. If a neighbor encoded its management address in character string format, you must configure the encoding format of the management address as string on the connecting port to guarantee normal communication with the neighbor.

To configure a management address to be advertised and its encoding format on a port or group of ports:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2/Layer 3 Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Allow LLDP to advertise the management address in LLDPDUs and configure the advertised management address.	lldp management-address-tlv [<i>ip-address</i>]	Optional. By default, the management address is sent through LLDPDUs. <ul style="list-style-type: none"> For a Layer 2 Ethernet port, the management address is the main IP address of the lowest-ID VLAN carried on the port. If none of the carried VLANs is assigned an IP address, no management address will be advertised. For a Layer 3 Ethernet port, the management address is its own IP address. If no IP address is configured for the Layer 3 Ethernet port, no management address will be advertised.

Step	Command	Remarks
4. Configure the encoding format of the management address as character string.	lldp management-address-format string	Optional. By default, the management address is encapsulated in the numeric format.

Setting other LLDP parameters

The Time to Live TLV carried in an LLDPDU determines how long the device information carried in the LLDPDU can be saved on a recipient device.

By setting the TTL multiplier, you can configure the TTL of locally sent LLDPDUs, which determines how long information about the local device can be saved on a neighboring device. The TTL is expressed by using the following formula:

$$\text{TTL} = \text{Min} (65535, (\text{TTL multiplier} \times \text{LLDPDU transmit interval}))$$

As the expression shows, the TTL can be up to 65535 seconds. TTLs greater than 65535 will be rounded down to 65535 seconds.

Configuration restrictions and guidelines

- To make sure that LLDP neighbors can receive LLDPDUs to update information about the current device before it ages out, configure both the LLDPDU transmit interval and delay to be less than the TTL.
- It is a good practice to set the LLDPDU transmit interval to be no less than four times the LLDPDU transmit delay.
- If the LLDPDU transmit delay is greater than the LLDPDU transmit interval, the device uses the LLDPDUs transmit delay as the transmit interval.

Configuration procedure

To change the TTL multiplier:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the TTL multiplier.	lldp hold-multiplier <i>value</i>	Optional. 4 by default.
3. Set the LLDPDU transmit interval.	lldp timer tx-interval <i>interval</i>	Optional. 30 seconds by default.
4. Set the LLDPDU transmit delay.	lldp timer tx-delay <i>delay</i>	Optional. 2 seconds by default.
5. Set the number of LLDPDUs sent each time fast LLDPDU transmission is triggered.	lldp fast-count <i>count</i>	Optional. 3 by default.

Setting an encapsulation format for LLDPDUs

LLDPDUs can be encapsulated in the following formats: Ethernet II or SNAP frames.

- With Ethernet II encapsulation configured, an LLDP port sends LLDPDUs in Ethernet II frames and processes only incoming, Ethernet II encapsulated LLDPDUs.
- With SNAP encapsulation configured, an LLDP port sends LLDPDUs in SNAP frames and processes only incoming, SNAP encapsulated LLDPDUs.

By default, Ethernet II frames encapsulate LLDPDUs. If the neighbor devices encapsulate LLDPDUs in SNAP frames, configure the encapsulation format for LLDPDUs as SNAP to guarantee normal communication with the neighbors.

To set the encapsulation format for LLDPDUs to SNAP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view or port group view.	<ul style="list-style-type: none"> • Enter Layer 2/Layer 3 Ethernet interface view: interface <i>interface-type interface-number</i> • Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Set the encapsulation format for LLDPDUs to SNAP.	lldp encapsulation snap	Ethernet II encapsulation format applies by default.

NOTE:

LLDP-CDP (Cisco Discovery Protocol) packets use only SNAP encapsulation.

Configuring CDP compatibility

To make your device work with Cisco IP phones, you must enable CDP compatibility.

If your LLDP-enabled device cannot recognize CDP packets, it does not respond to the requests of Cisco IP phones for the voice VLAN ID configured on the device. As a result, a requesting Cisco IP phone sends voice traffic without any tag to your device, and, as a result, your device cannot differentiate the voice traffic from other types of traffic.

With CDP compatibility enabled, your device can receive and recognize CDP packets from a Cisco IP phone and respond with CDP packets, which carry the voice VLAN configuration TLVs. According to the voice VLAN configuration TLVs, the IP phone automatically configures the voice VLAN. As a result, the voice traffic is confined in the configured voice VLAN, and differentiated from other types of traffic.

For more information about voice VLANs, see "[Configuring a voice VLAN](#)."

Configuration prerequisites

Before you configure CDP compatibility, complete the following tasks:

- Globally enable LLDP.
- Enable LLDP on the port connecting to an IP phone and configure the port to operate in TxRx mode.

Configuring CDP compatibility

CAUTION:

The maximum TTL value that CDP allows is 255 seconds. To make CDP-compatible LLDP work properly with Cisco IP phones, be sure that the product of the TTL multiplier and the LLDPDU transmit interval is less than 255 seconds.

CDP-compatible LLDP operates in one of the follows modes:

- **TxRx**—CDP packets can be transmitted and received.
- **Disable**—CDP packets cannot be transmitted or received.

To make CDP-compatible LLDP take effect on specific ports, first enable CDP-compatible LLDP globally, and then configure CDP-compatible LLDP to operate in TxRx mode.

To enable LLDP to be compatible with CDP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable CDP compatibility globally.	lldp compliance cdp	Disabled by default.
3. Enter Ethernet interface view or port group view.	<ul style="list-style-type: none">• Enter Layer 2/Layer 3 Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i>• Enter port group view: port-group manual <i>port-group-name</i>	Use either command.
4. Configure CDP-compatible LLDP to operate in TxRx mode.	lldp compliance admin-status cdp txrx	Disable mode by default.

Enabling LLDP to automatically discover IP phones

In a traditional voice VLAN network, the switch maps the source MAC addresses of IP phones to a limited number of OUI addresses to allow them to access the network. This method restricts the types of IP phones on the network, if the IP phones with the source MAC addresses match the same OUI address are categorized as a type.

To break the restriction, you can enable the switch to automatically discover IP phones through LLDP. With this function, the switch can automatically discover the peer, and exchange LLDP TLVs with the peer. If the LLDP System Capabilities TLV received on a port shows that the peer is phone capable, the switch determines that the peer is an IP phone and sends an LLDP TLV carrying the voice VLAN configuration to the peer.

When the IP phone discovery process is complete, the port will automatically join the voice VLAN and improve the transmission priority of the voice traffic for the IP phone. To ensure that the IP phone can pass authentication, the switch will add the MAC address of the IP phone to the MAC address table.

For more information about voice VLANs, see "[Configuring a voice VLAN.](#)"

Configuration prerequisites

Before you enable the switch to automatically discover IP phones through LLDP, complete the following tasks:

- Enable LLDP globally and on ports.
- Configure voice VLANs.

Configuration procedure

To enable LLDP to automatically discover IP phones:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable LLDP to automatically discover IP phones.	voice vlan track lldp	Disabled by default.

ⓘ IMPORTANT:

- When the switch is enabled to automatically discover IP phones through LLDP, you can connect at most five IP phones to each port of the switch.
- You cannot use this function together with CDP compatibility.

Configuring LLDP to advertise a specific voice VLAN

Voice VLAN advertisement through LLDP is available only for LLDP-enabled IP phones. If CDP-compatibility is enabled, this feature is also available for CDP-enabled IP phones. For more information about CDP compatibility, see "[Configuring CDP compatibility](#)." For more information about the voice VLANs, see "[Configuring a voice VLAN](#)."

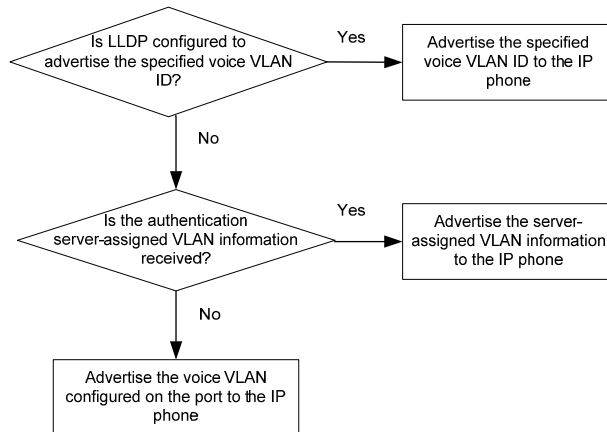
Configuration guidelines

Use this feature in one of the following scenarios:

- Decrease the voice VLAN processing delay in an IRF fabric.
On an LLDP-enabled port, LLDP advertises the voice VLAN information to the IP phone connected to the port. When a packet arrives on the port, the switch compares the source MAC address against its voice device OUI list. If a match is found, the switch learns the MAC address in the voice VLAN, and promotes the forwarding priority for the packet. Because this process is completed in software, in an IRF fabric, MAC address learning and synchronization of the learned MAC address entry to all member devices introduces an undesirable delay. Directly specifying the voice VLAN to be advertised by LLDP enables the IRF fabric to learn and synchronize MAC address entries faster in hardware.
- Avoid configuring the voice VLAN function on a port.

[Figure 75](#) shows the procedure of voice VLAN advertisement through LLDP.

Figure 75 Voice VLAN advertisement through LLDP



With the received voice VLAN information, the IP phone automatically completes the voice VLAN configuration, including the voice VLAN ID, tagging status, and priority. This voice VLAN can be the voice VLAN directly specified for LLDP advertisement, the voice VLAN configured on the port, or the voice VLAN assigned by a server, depending on your configuration.

To identify the voice VLAN advertised by LLDP, execute the **display lldp local-information** command, and examine the MED information fields in the command output.

Configuration procedure

To configure LLDP to advertise a specific voice VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable LLDP globally.	lldp enable	By default, LLDP is globally enabled.
3. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use one of the commands.
4. Enable LLDP.	lldp enable	Optional. By default, LLDP is enabled on a port.
5. Configure LLDP to advertise a specific voice VLAN.	lldp voice-vlan <i>vlan-id</i>	By default, LLDP advertises the voice VLAN configured on the port.

Dynamically advertising server-assigned VLANs through LLDP

Dynamic advertisement of server-assigned VLANs through LLDP must work with 802.1X or MAC authentication, and is available only for LLDP-enabled IP phones. If 802.1X authentication is used, make sure the IP phones also support 802.1X authentication.

To implement this function for an IP phone, perform the following configuration tasks:

- Enable LLDP globally and on the port connected to the IP phone.
- Configure 802.1X or MAC authentication to make sure the IP phone can pass security authentication. For more information about 802.1X authentication, MAC authentication, and VLAN assignment by servers, see *Security Configuration Guide*.
- Configure VLAN authorization for the IP phone on the authentication server.

After the IP phone passes authentication, LLDP advertises the server-assigned VLAN in the Network Policy TLV to the IP phone. The IP phone will send its traffic tagged with the assigned VLAN.

Configuring LLDP trapping

LLDP trapping notifies the network management system (NMS) of events such as newly-detected neighboring devices and link malfunctions.

LLDP traps are sent periodically, and the interval is configurable. To prevent excessive LLDP traps from being sent when the topology is unstable, set a trap transmit interval for LLDP.

To configure LLDP trapping:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view or port group view.	<ul style="list-style-type: none">• Enter Layer 2/Layer 3 Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i>• Enter port group view: port-group manual <i>port-group-name</i>	Use either command.
3. Enable LLDP trapping.	lldp notification remote-change enable	Disabled by default.
4. Return to system view.	quit	N/A
5. Set the LLDP trap transmit interval.	lldp timer notification-interval <i>interval</i>	Optional. 5 seconds by default.

Displaying and maintaining LLDP

Task	Command	Remarks
Display the global LLDP information or the information contained in the LLDP TLVs to be sent through a port.	display lldp local-information [global interface <i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the information contained in the LLDP TLVs sent from neighboring devices.	display lldp neighbor-information [brief interface <i>interface-type interface-number</i> [brief] list [system-name <i>system-name</i>]] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display LLDP statistics.	display lldp statistics [global interface <i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display LLDP status of a port.	display lldp status [interface <i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display types of advertisable optional LLDP TLVs.	display lldp tlv-config [interface <i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

LLDP configuration examples

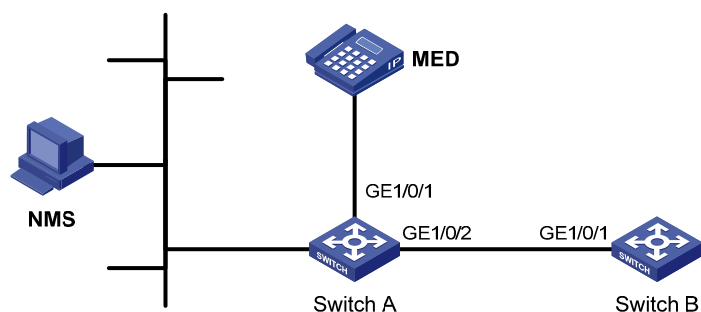
Basic LLDP configuration example

Network requirements

As shown in [Figure 76](#), the NMS and Switch A are located in the same Ethernet. An MED device and Switch B are connected to GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 of Switch A.

Enable LLDP on the ports of Switch A and Switch B to monitor the link between Switch A and Switch B and the link between Switch A and the MED device on the NMS.

Figure 76 Network diagram



Configuration procedure

1. Configure Switch A:


```
# Enable LLDP globally.
<SwitchA> system-view
[SwitchA] lldp enable
```

Enable LLDP on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2. (You can skip this step because LLDP is enabled on ports by default.) Set the LLDP operating mode to Rx.

```
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] lldp enable
[SwitchA-GigabitEthernet1/0/1] lldp admin-status rx
[SwitchA-GigabitEthernet1/0/1] quit
[SwitchA] interface gigabitethernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] lldp enable
[SwitchA-GigabitEthernet1/0/2] lldp admin-status rx
[SwitchA-GigabitEthernet1/0/2] quit
```

2. Configure Switch B:

Enable LLDP globally.

```
<SwitchB> system-view
[SwitchB] lldp enable
```

Enable LLDP on GigabitEthernet1/0/1. (You can skip this step because LLDP is enabled on ports by default.) Set the LLDP operating mode to Tx.

```
[SwitchB] interface gigabitethernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] lldp enable
[SwitchB-GigabitEthernet1/0/1] lldp admin-status tx
[SwitchB-GigabitEthernet1/0/1] quit
```

3. Verify the configuration:

Display the global LLDP status and port LLDP status on Switch A.

```
[SwitchA] display lldp status
Global status of LLDP: Enable
The current number of LLDP neighbors: 2
The current number of CDP neighbors: 0
LLDP neighbor information last changed time: 0 days,0 hours,4 minutes,40 seconds
Transmit interval          : 30s
Hold multiplier            : 4
Reinit delay               : 2s
Transmit delay             : 2s
Trap interval              : 5s
Fast start times           : 3
```

```
Port 1 [GigabitEthernet1/0/1]:
Port status of LLDP        : Enable
Admin status                : Rx_Only
Trap flag                   : No
Polling interval           : 0s
```

```
Number of neighbors: 1
Number of MED neighbors : 1
Number of CDP neighbors : 0
Number of sent optional TLV : 0
Number of received unknown TLV : 0
```

```
Port 2 [GigabitEthernet1/0/2]:
```



```
Port status of LLDP          : Enable
Admin status                 : Rx_Only
Trap flag                    : No
Polling interval             : 0s
```

```
Number of neighbors:       1
Number of MED neighbors    : 0
Number of CDP neighbors    : 0
Number of sent optional TLV : 0
Number of received unknown TLV : 3
```

As the sample output shows, GigabitEthernet 1/0/1 of Switch A connects to an MED device, and GigabitEthernet 1/0/2 of Switch A connects to a non-MED device. Both ports operate in Rx mode, and they only receive LLDPDUs.

Remove the link between Switch A and Switch B and then display the global LLDP status and port LLDP status on Switch A.

```
[SwitchA] display lldp status
Global status of LLDP: Enable
The current number of LLDP neighbors: 1
The current number of CDP neighbors: 0
LLDP neighbor information last changed time: 0 days,0 hours,5 minutes,20 seconds
Transmit interval          : 30s
Hold multiplier            : 4
Reinit delay               : 2s
Transmit delay             : 2s
Trap interval              : 5s
Fast start times           : 3
```

```
Port 1 [GigabitEthernet1/0/1]:
Port status of LLDP          : Enable
Admin status                 : Rx_Only
Trap flag                    : No
Polling interval             : 0s
```

```
Number of neighbors        : 1
Number of MED neighbors    : 1
Number of CDP neighbors    : 0
Number of sent optional TLV : 0
Number of received unknown TLV : 5
```

```
Port 2 [GigabitEthernet1/0/2]:
Port status of LLDP          : Enable
Admin status                 : Rx_Only
Trap flag                    : No
Polling interval             : 0s
```

```
Number of neighbors        : 0
Number of MED neighbors    : 0
Number of CDP neighbors    : 0
```

```
Number of sent optional TLV      : 0
Number of received unknown TLV  : 0
```

As the sample output shows, GigabitEthernet 1/0/2 of Switch A does not connect to any neighboring devices.

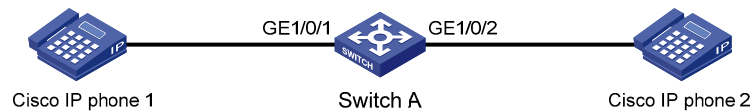
CDP-compatible LLDP configuration example

Network requirements

As shown in [Figure 77](#), GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 of Switch A are each connected to a Cisco IP phone. The two IP phones send out tagged voice traffic.

Configure voice VLAN 2 on Switch A. Enable CDP compatibility of LLDP on Switch A to allow the Cisco IP phones to automatically configure the voice VLAN, confining their voice traffic within the voice VLAN and isolating the voice traffic from other types of traffic.

Figure 77 Network diagram



Configuration procedure

1. Configure a voice VLAN on Switch A:

```
# Create VLAN 2.
```

```
<SwitchA> system-view
[SwitchA] vlan 2
[SwitchA-vlan2] quit
```

```
# Set the link type of GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 to trunk and enable voice VLAN on them.
```

```
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] port link-type trunk
[SwitchA-GigabitEthernet1/0/1] voice vlan 2 enable
[SwitchA-GigabitEthernet1/0/1] quit
[SwitchA] interface gigabitethernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] port link-type trunk
[SwitchA-GigabitEthernet1/0/2] voice vlan 2 enable
[SwitchA-GigabitEthernet1/0/2] quit
```

2. Configure CDP-compatible LLDP on Switch A:

```
# Enable LLDP globally and enable LLDP to be compatible with CDP globally.
```

```
[SwitchA] lldp enable
[SwitchA] lldp compliance cdp
```

```
# Enable LLDP (you can skip this step because LLDP is enabled on ports by default.), configure LLDP to operate in TxRx mode, and configure CDP-compatible LLDP to operate in TxRx mode on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2.
```

```
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] lldp enable
[SwitchA-GigabitEthernet1/0/1] lldp admin-status txrx
[SwitchA-GigabitEthernet1/0/1] lldp compliance admin-status cdp txrx
```

```
[SwitchA-GigabitEthernet1/0/1] quit
[SwitchA] interface gigabitethernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] lldp enable
[SwitchA-GigabitEthernet1/0/2] lldp admin-status txrx
[SwitchA-GigabitEthernet1/0/2] lldp compliance admin-status cdp txrx
[SwitchA-GigabitEthernet1/0/2] quit
```

3. Verify the configuration:

Display the neighbor information on Switch A.

```
[SwitchA] display lldp neighbor-information
```

```
CDP neighbor-information of port 1[GigabitEthernet1/0/1]:
```

```
  CDP neighbor index : 1
  Chassis ID         : SEP00141CBCDBFF
  Port ID            : Port 1
  Software version   : P0030301MFG2
  Platform           : Cisco IP Phone 7960
  Duplex             : Full
```

```
CDP neighbor-information of port 2[GigabitEthernet1/0/2]:
```

```
  CDP neighbor index : 2
  Chassis ID         : SEP00141CBCDBFF
  Port ID            : Port 1
  Software version   : P0030301MFG2
  Platform           : Cisco IP Phone 7960
  Duplex             : Full
```

As the sample output shows, Switch A has discovered the IP phones connected to GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, and has obtained their LLDP device information.

Configuring a service loopback group (available only on the 5500 EI)

Overview

To increase traffic redirecting throughput, you can bundle multiple Ethernet ports of a device together to increase bandwidth and implement load sharing. These ports that act as a logical link form a service loopback group. A service loopback group must contain at least one Ethernet port as its member port, called a service loopback port.

For example, by assigning three Ethernet ports of the same device to a service loopback group, you can create a logical link whose bandwidth can be as high as the total bandwidth of these three ports. In addition, service traffic is load balanced among these ports.

Service types of service loopback groups

A service loopback group is applicable to a specified service type. The service types supported by the service loopback group are:

- **Tunnel**—Supports unicast tunnel traffic

Requirements on service loopback ports

Before you assign a port to a service loopback group, ensure the following:

- The port supports the services type or types of the service loopback group.
- The port is not configured with multiple spanning tree protocol (MSTP), Link Layer Discovery Protocol (LLDP), Neighbor Discovery Protocol (NDP), 802.1X, MAC address authentication, port security mode, or IP source guard, or as the member port of an isolation group.
- The link type of the port is access.
- The port is not a member of any Ethernet link aggregation group or service loopback group.

States of service loopback ports

A member port in a service loopback group is a service loopback port, which can be in either of the following states:

- **Selected**—A selected port can loop back user traffic.
- **Unselected**—An unselected port cannot loop back user traffic.

The number of selected ports is limited in a service loopback group.

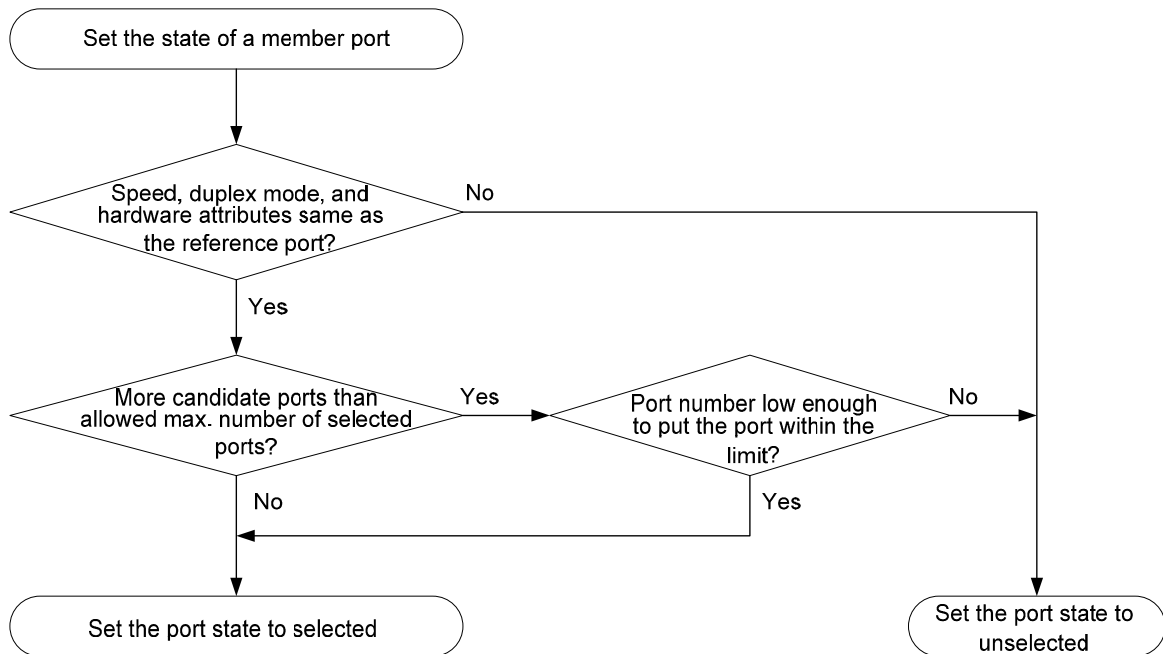
Setting the state of service loopback ports

The system sets the state of each member port in a service loopback group to selected or unselected by using the following workflow:

1. Select the full-duplex port with the highest rate as the reference port. If two ports have the same duplex mode and speed, the one with the lower port number wins.

- Set the state of each member port in the service loopback group.

Figure 78 Setting the state of each member port in a service loopback group



NOTE:

Every time a new port is assigned to a service loopback group, the system resets the state of the member ports in the service loopback group according to the process described previously.

Configuration restrictions and guidelines

- A service loopback group may be referenced by multiple features once it is created, and can process service traffic only after it is referenced.
- You can change the service type of an existing service loopback group. For the change to be successful, make sure that the service group has not been referenced, the attributes of all member ports (if any) do not conflict with the target service type, and no service loopback group has been created for the target service type, because only one service loopback group is allowed for a service type.
- You cannot remove a service loopback group that other features reference.

Configuration procedure

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a service loopback group and specify its service type.	service-loopback group <i>number</i> type tunnel	N/A

Step	Command	Remarks
3. Enter Layer 2 Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Assign the Ethernet interface to the specified service loopback group.	port service-loopback group <i>number</i>	By default, a port does not belong to any service loopback group. You can perform this command on different ports to assign multiple ports to a service loopback group.

Displaying and maintaining service loopback groups

Task	Command	Remarks
Display information about the specified service loopback group or all service loopback groups.	display service-loopback group [<i>number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Service loopback group configuration example

Network requirements

All the Ethernet ports of Device A support the tunnel service. Assign GigabitEthernet1/0/1 through GigabitEthernet1/0/3 to a service loopback group to increase bandwidth and achieve load sharing for tunnel traffic.

Configuration procedure

```
# Create service loopback group 1, and specify the service type as tunnel (unicast tunnel service).
<DeviceA> system-view
[DeviceA] service-loopback group 1 type tunnel

# Disable MSTP, LLDP, and NDP on GigabitEthernet1/0/1 through GigabitEthernet1/0/3 and then
assign them to service loopback group 1.
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] undo stp enable
[DeviceA-GigabitEthernet1/0/1] undo lldp enable
[DeviceA-GigabitEthernet1/0/1] undo ndp enable
[DeviceA-GigabitEthernet1/0/1] port service-loopback group 1
[DeviceA-GigabitEthernet1/0/1] quit
[DeviceA] interface gigabitethernet 1/0/2
[DeviceA-GigabitEthernet1/0/2] undo stp enable
[DeviceA-GigabitEthernet1/0/2] undo lldp enable
[DeviceA-GigabitEthernet1/0/2] undo ndp enable
[DeviceA-GigabitEthernet1/0/2] port service-loopback group 1
[DeviceA-GigabitEthernet1/0/2] quit
```

```
[DeviceA] interface gigabitethernet 1/0/3
[DeviceA-GigabitEthernet1/0/3] undo stp enable
[DeviceA-GigabitEthernet1/0/3] undo lldp enable
[DeviceA-GigabitEthernet1/0/3] undo ndp enable
[DeviceA-GigabitEthernet1/0/3] port service-loopback group 1
[DeviceA-GigabitEthernet1/0/3] quit

# Create logical interface Tunnel 1 and reference service loopback group 1 on Tunnel 1.
[DeviceA] interface tunnel 1
[DeviceA-Tunnel1] service-loopback-group 1
```

Configuring MVRP

Overview

Multiple Registration Protocol (MRP) is an attribute registration protocol and transmits attribute messages. Multiple VLAN Registration Protocol (MVRP) is a typical MRP application. MVRP propagates and learns VLAN configuration among devices. MVRP enables a device to propagate the local VLAN configuration to the other devices, receive VLAN configuration from other devices, and dynamically update the local VLAN configuration (including the active VLANs and the ports through which a VLAN can be reached). MVRP makes sure that all MVRP-enabled devices in a LAN maintain the same VLAN configuration, and reduces the VLAN configuration workload. When the network topology changes, MVRP can propagate and learn VLAN configuration information again according to the new topology, and real-time synchronize the network topology.

MRP is an enhanced version of Generic Attribute Registration Protocol (GARP) and improves the declaration efficiency. MVRP is an enhanced version of GARP VLAN Registration Protocol (GVRP). MVRP delivers the following benefits over GVRP:

- GVRP does not support the multiple spanning tree instance (MSTI). MVRP runs on a per-MSTI basis, and implements per-VLAN redundant link calculation and load sharing.
- MVRP decreases the number of packets transmitted for the same amount of VLAN configuration, and improves the declaration efficiency.

For more information about GVRP, see "[Configuring GVRP](#)." For more information about MSTI, see "[Configuring spanning tree protocols](#)."

Introduction to MRP

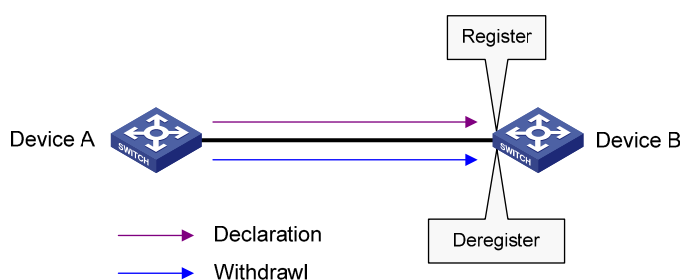
MRP allows participants in the same LAN to declare, propagate, and register information (for example, VLAN information) on a per Multiple Spanning Tree Instance (MSTI) basis.

MRP implementation

Each port that participates in an MRP application (for example, MVRP) is called an "MRP participant". Similarly, a port that participates in an MVRP application is called an "MVRP participant."

As shown in [Figure 79](#), an MRP participant registers and deregisters its attribute values on other MRP participants by sending declarations and withdrawals, and registers and deregisters the attribute values of other participants according to the received declarations and withdrawals. MRP rapidly propagates the configuration information of an MRP participant throughout the LAN.

Figure 79 MRP implementation



MVRP registers and deregisters VLAN attributes as follows:

- When a port receives the declaration of a VLAN attribute, the port registers the VLAN and joins the VLAN.
- When a port receives the withdrawal of a VLAN attribute, the port deregisters the VLAN and leaves the VLAN.

Figure 79 shows a simple MVRP implementation on an MSTI. In a network with multiple MSTIs, VLAN registration and deregistration are performed on a per-MSTI basis.

MRP messages

MRP exchanges information among MRP participants by advertising MRP messages, including Join, New, Leave, and LeaveAll. Join and New messages are declarations, and Leave and LeaveAll messages are withdrawals.

- Join message
 - An MRP participant sends Join messages when it wishes to declare the attribute values configured on it and receives Join messages from other MRP participants.
 - When receiving a Join message, an MRP participant sends a Join message to all participants except the sender.

Join messages fall into the following types:

- **JoinEmpty**—An MRP participant sends JoinEmpty messages to declare attribute values that it has not registered. For example, when a static VLAN exists on a device, the attribute of the VLAN on the device is not changed even if the device learns the VLAN again through MRP. In this case, the Join message for the VLAN attribute is a JoinEmpty message, because the VLAN attribute is not registered.
- **JoinIn**—An MRP participant sends JoinIn messages to declare attribute values that it has registered. For example, when the device learns a VLAN through MRP messages, and dynamically creates the VLAN, the Join message for the VLAN attribute is a JoinIn message.
- New message

Similar to a Join message, a New message enables MRP participants to register attributes.

 - When the Multiple Spanning Tree Protocol (MSTP) topology changes, an MRP participant sends New messages to declare the topology change.
 - On receiving a New message, an MRP participant sends a New message out of each port except the receiving port.
- Leave message
 - An MRP participant sends Leave messages when it wishes other participants to deregister the attributes that it has deregistered.
 - When receiving a Leave message, an MRP participant sends a Leave message to all participants except the sender.
- LeaveAll message
 - Each MRP participant is configured with an individual LeaveAll timer. When the timer expires, the MRP participant sends LeaveAll messages to the remote participants, so that the local participant and the remote participants deregister all attributes and re-register all attributes. This process periodically clears the useless attributes in the network.
 - On receiving a LeaveAll message, MRP determines whether to send a Join message to request the sender to re-register these attributes according to attribute status.

MRP timers

The implementation of MRP uses the following timers to control MRP message transmission.

- Periodic timer

On startup, an MRP participant starts its own Periodic timer to control MRP message transmission. The MRP participant collects the MRP messages to be sent before the Periodic timer expires, and sends the MRP messages in as few packets as possible when the Periodic timer expires and meanwhile restarts the Periodic timer. This mechanism reduces the number of MRP protocol packets periodically sent.

You can enable or disable the Periodic timer at the CLI. When you disable the Periodic timer, MRP will not periodically send MRP messages.

- Join timer

The Join timer controls the transmission of Join messages. To make sure that Join messages can be reliably transmitted to other participants, an MRP participant waits for a period of the Join timer after sending a Join message. If the participant receives JoinIn messages from other participants and the attributes in the JoinIn messages are the same as the sent Join messages before the Join timer expires, the participant does not re-send the Join message. When both the Join timer and the Periodic timer expire, the participant re-sends the Join message.

- Leave timer

The Leave timer controls the deregistration of attributes. When an MRP participant wishes other participants to deregister its attributes, it sends a Leave message. On receiving a Leave message, MRP starts the Leave timer, and deregisters the attributes if it does not receive any Join message for the attributes before the Leave timer expires. When an MRP participant sends or receives LeaveAll messages, it starts the Leave timer. MRP deregisters the attributes in the LeaveAll messages if it does not receive any Join message for the attributes before the Leave timer expires.

- LeaveAll timer

On startup, an MRP participant starts its own LeaveAll timer. When the LeaveAll timer expires, MRP sends out a LeaveAll message and restarts the LeaveAll timer. On receiving the LeaveAll message, other participants re-register all the attributes and re-start their LeaveAll timer.

When the LeaveAll timer of an MRP participant expires, the MRP participant sends LeaveAll messages to the remote participants. On receiving a LeaveAll message, a remote participant restarts its LeaveAll timer, and stops sending out LeaveAll messages. This mechanism effectively reduces the number of LeaveAll messages in the network.

To avoid the case that the LeaveAll timer of a fixed participant always first expires, the switch randomly changes the LeaveAll timer within a certain range when the MRP participant restarts its LeaveAll timer.

MVRP registration modes

The VLAN information propagated by MVRP includes not only locally, manually configured static VLAN information but also dynamic VLAN information from other devices.

VLANs created manually, locally are called "static VLANs", and VLANs learned through MVRP are called "dynamic VLANs". The following MVRP registration modes are available.

- Normal

An MVRP participant in normal registration mode performs dynamic VLAN registrations and deregistrations, and sends declarations and withdrawals for dynamic and static VLANs.

- Fixed

An MVRP participant in fixed registration mode disables deregistering dynamic VLANs, sends declarations for dynamic VLANs and static VLANs, and drops received MVRP protocol packets. As a result, an MVRP participant port in fixed registration mode does not deregister or register dynamic VLANs.

- Forbidden

An MVRP participant in forbidden registration mode disables registering dynamic VLANs, sends declarations for dynamic VLANs and static VLANs, and drops received MVRP protocol packets. As a result, an MVRP participant in forbidden registration mode does not register dynamic VLANs, and does not re-register a dynamic VLAN when the VLAN is deregistered.

Protocols and standards

IEEE 802.1ak *IEEE Standard for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks – Amendment 07: Multiple Registration Protocol*

MVRP configuration task list

Task	Remarks
Configuration prerequisites	Required.
Enabling MVRP	Required.
Configuring the MVRP registration mode	Optional.
Configuring MRP timers	Optional.
Enabling GVRP compatibility	Optional.

Configuration prerequisites

Before configuring MVRP, perform the following tasks:

- Make sure that all MSTIs in the network are effective and each MSTI is mapped to an existing VLAN on each device in the network, because MVRP runs on a per-MSTI basis.
- Configure the involved ports as trunk ports, because MVRP is available only on trunk ports.

Enabling MVRP

Configuration restrictions and guidelines

- MVRP can work with STP, RSTP, or MSTP, but not other link layer topology protocols, including service loopback, PVST, RRPP, and Smart Link. Ports blocked by STP, RSTP, or MSTP can receive and send MVRP protocol packets. For more information about service loopback, see "[Configuring a service loopback group \(available only on the 5500 EI\)](#)." For more information about STP, RSTP, MSTP, and PVST, see "[Configuring spanning tree protocols](#)." For more information about RRPP and Smart Link, see *High Availability Configuration Guide*.
- Do not enable both MVRP and remote port mirroring on a port. Otherwise, MVRP may register the remote probe VLAN to incorrect ports, which would cause the monitor port to receive undesired

duplicates. For more information about port mirroring, see *Network Management and Monitoring Configuration Guide*.

- Enabling MVRP on a Layer 2 aggregate interface enables both the aggregate interface and all Selected member ports in the link aggregation group to participate in dynamic VLAN registration and deregistration.

Configuration procedure

To enable MVRP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable MVRP globally.	mvrp global enable	By default, MVRP is globally disabled. To enable MVRP on a port, first enable MVRP globally.
3. Enter interface view.	<ul style="list-style-type: none"> • Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> • Enter port group view: port-group manual <i>port-group-name</i> 	Use one of the commands.
4. Configure the port to permit the specified VLANs.	port trunk permit vlan { <i>vlan-list</i> all }	By default, a trunk port permits only VLAN 1. Make sure that the trunk port permits all registered VLANs. For more information about the port trunk permit vlan { <i>vlan-list</i> all } command, see <i>Layer 2—LAN Switching Command Reference</i> .
5. Enable MVRP on the port.	mvrp enable	By default, MVRP is disabled on a port.

Configuring the MVRP registration mode

To configure the MVRP registration mode:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter interface view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use one of the commands.
3. Configure the MVRP registration mode.	mvrp registration { fixed forbidden normal }	Optional. The default setting is normal registration mode.

Configuring MRP timers

⚠ CAUTION:

The MRP timers apply to all MRP applications, for example, MVRP, on a port. To avoid frequent VLAN registrations and deregistrations, use the same MRP timers throughout the network.

Each port maintains its own Periodic, Join, and LeaveAll timers, and each attribute of a port maintains a Leave timer.

To configure MRP timers:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use one of the commands.
3. Configure the LeaveAll timer.	mrp timer leaveall <i>timer-value</i>	Optional. The default setting is 1000 centiseconds.
4. Configure the Join timer.	mrp timer join <i>timer-value</i>	Optional. The default setting is 20 centiseconds.
5. Configure the Leave timer.	mrp timer leave <i>timer-value</i>	Optional. The default setting is 60 centiseconds.
6. Configure the Periodic timer.	mrp timer periodic <i>timer-value</i>	Optional. The default setting is 100 centiseconds.

Table 28 shows the value ranges for Join, Leave, and LeaveAll timers and their dependencies.

- If you set a timer to a value beyond the allowed value range, your configuration will fail. To do that, you can change the allowed value range by tuning the value of another related timer.
- To restore the default settings of the timers, restore the Join timer first, followed by the Leave and LeaveAll timers.

Table 28 Dependencies of the Join, Leave, and LeaveAll timers

Timer	Lower limit	Upper limit
Join	20 centiseconds	Half the Leave timer
Leave	Twice the Join timer	LeaveAll timer
LeaveAll	Leave timer on each port	32760 centiseconds

NOTE:

You can restore the Periodic timer to the default at any time.

Enabling GVRP compatibility

MVRP can be compatible with GVRP. When the peer device supports GVRP, you can enable GVRP compatibility on the local end, so that the local end can receive and send MVRP and GVRP protocol packets at the same time.

Configuration restrictions and guidelines

- MVRP with GVRP compatibility enabled can work together with STP or RSTP, but cannot work together with MSTP. When MVRP with GVRP compatibility enabled works with MSTP, the network might operate improperly.
- When GVRP compatibility is enabled for MVRP, HP recommends disabling the Period timer. Otherwise, the VLAN status might frequently change when the system is busy.

Configuration procedure

To enable GVRP compatibility:

Step	Command	Remarks
1. Enter system view	system-view	N/A
2. Enable GVRP compatibility	mvrp gvrp-compliance enable	By default, GVRP compatibility is disabled.

Displaying and maintaining MVRP

Task	Command	Remarks
Display the MVRP status of the specified port and each MVRP interface in the specified VLAN.	display mvrp state interface <i>interface-type interface-number</i> vlan <i>vlan-id</i> [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the MVRP running status.	display mvrp running-status [interface <i>interface-list</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the MVRP statistics.	display mvrp statistics [interface <i>interface-list</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the dynamic VLAN operation information of the specified port.	display mvrp vlan-operation interface <i>interface-type interface-number</i> [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear the MVRP statistics of the specified ports.	reset mvrp statistics [interface <i>interface-list</i>]	Available in user view

Configuration example for MVRP in normal registration mode

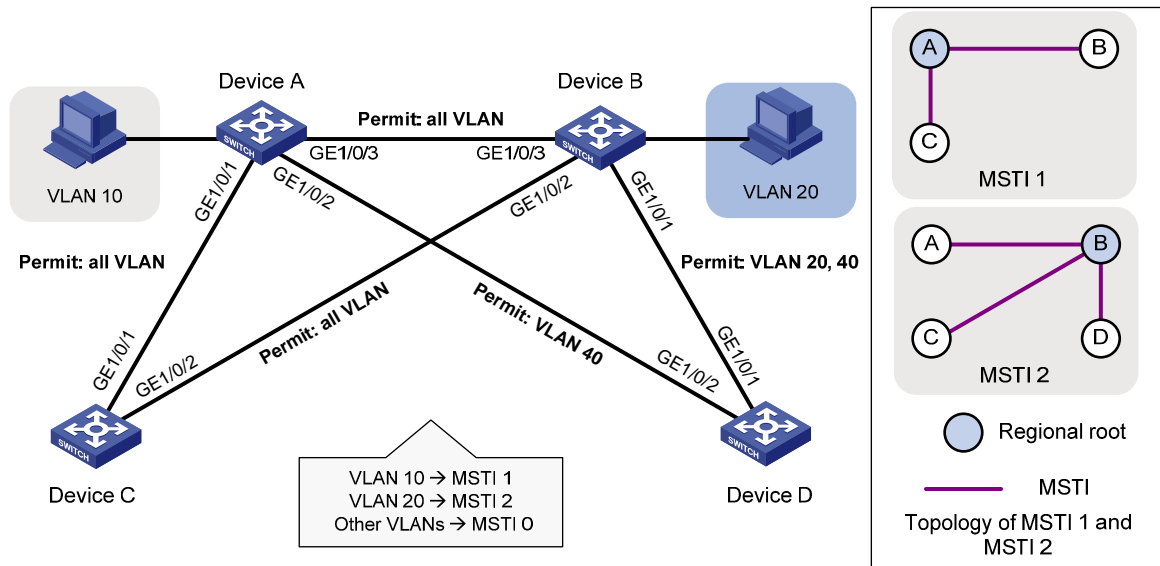
Network requirements

As shown in [Figure 80](#), configure MSTP, map VLAN 10 to MSTI 1, map VLAN 20 MST 2, and map the other VLANs to MSTI 0.

Configure MVRP and set the MVRP registration mode to normal, so that Device A, Device B, Device C, and Device D can register and deregister dynamic and static VLANs and keep identical VLAN configuration for each MSTI.

When the network is stable, set the MVRP registration mode to fixed on the port that connecting Device B to Device A, so that the dynamic VLANs on Device B are not de-registered.

Figure 80 Network diagram



Configuration procedure

Configuring Device A

```
# Enter MST region view.
<DeviceA> system-view
[DeviceA] stp region-configuration

# Configure the MST region name, VLAN-to-instance mappings, and revision level.
[DeviceA-mst-region] region-name example
[DeviceA-mst-region] instance 1 vlan 10
[DeviceA-mst-region] instance 2 vlan 20
[DeviceA-mst-region] revision-level 0

# Manually activate the MST region configuration.
[DeviceA-mst-region] active region-configuration
[DeviceA-mst-region] quit

# Configure Device A as the primary root bridge of MSTI 1.
[DeviceA] stp instance 1 root primary

# Globally enable the spanning tree feature.
[DeviceA] stp enable

# Globally enable MVRP.
[DeviceA] mvrp global enable

# Configure port GigabitEthernet 1/0/1 as a trunk port, and configure it to permit all VLANs.
[DeviceA] interface gigabitEthernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] port link-type trunk
[DeviceA-GigabitEthernet1/0/1] port trunk permit vlan all

# Enable MVRP on port GigabitEthernet 1/0/1.
[DeviceA-GigabitEthernet1/0/1] mvrp enable
[DeviceA-GigabitEthernet1/0/1] quit
```



```

# Configure port GigabitEthernet1/0/2 as a trunk port, and configure it to permit VLAN 40.
[DeviceA] interface gigabitethernet 1/0/2
[DeviceA-GigabitEthernet1/0/2] port link-type trunk
[DeviceA-GigabitEthernet1/0/2] port trunk permit vlan 40

# Enable MVRP on port GigabitEthernet1/0/2.
[DeviceA-GigabitEthernet1/0/2] mvrp enable
[DeviceA-GigabitEthernet1/0/2] quit

# Configure port GigabitEthernet 1/0/3 as a trunk port, and configure it to permit all VLANs.
[DeviceA] interface gigabitethernet 1/0/3
[DeviceA-GigabitEthernet1/0/3] port link-type trunk
[DeviceA-GigabitEthernet1/0/3] port trunk permit vlan all

# Enable MVRP on port GigabitEthernet 1/0/3.
[DeviceA-GigabitEthernet1/0/3] mvrp enable
[DeviceA-GigabitEthernet1/0/3] quit

# Create VLAN 10.
[DeviceA] vlan 10
[DeviceA-vlan10] quit

```

Configuring Device B

```

# Enter MST region view.
<DeviceB> system-view
[DeviceB] stp region-configuration

# Configure the MST region name, VLAN-to-instance mappings, and revision level.
[DeviceB-mst-region] region-name example
[DeviceB-mst-region] instance 1 vlan 10
[DeviceB-mst-region] instance 2 vlan 20
[DeviceB-mst-region] revision-level 0

# Manually activate the MST region configuration.
[DeviceB-mst-region] active region-configuration
[DeviceB-mst-region] quit

# Configure Device B as the primary root bridge of MSTI 2.
[DeviceB] stp instance 2 root primary

# Globally enable the spanning tree feature.
[DeviceB] stp enable

# Globally enable MVRP.
[DeviceB] mvrp global enable

# Configure port GigabitEthernet 1/0/1 as a trunk port, and configure it to permit VLANs 20 and 40.
[DeviceB] interface gigabitethernet 1/0/1
[DeviceB-GigabitEthernet1/0/1] port link-type trunk
[DeviceB-GigabitEthernet1/0/1] port trunk permit vlan 20 40

# Enable MVRP on port GigabitEthernet 1/0/1.
[DeviceB-GigabitEthernet1/0/1] mvrp enable
[DeviceB-GigabitEthernet1/0/1] quit

# Configure port GigabitEthernet1/0/2 as a trunk port, and configure it to permit all VLANs.

```

```

[DeviceB] interface gigabitethernet 1/0/2
[DeviceB-GigabitEthernet1/0/2] port link-type trunk
[DeviceB-GigabitEthernet1/0/2] port trunk permit vlan all
# Enable MVRP on port GigabitEthernet1/0/2.
[DeviceB-GigabitEthernet1/0/2] mvrp enable
[DeviceB-GigabitEthernet1/0/2] quit
# Configure port GigabitEthernet 1/0/3 as a trunk port, and configure it to permit all VLANs.
[DeviceB] interface gigabitethernet 1/0/3
[DeviceB-GigabitEthernet1/0/3] port link-type trunk
[DeviceB-GigabitEthernet1/0/3] port trunk permit vlan all
# Enable MVRP on port GigabitEthernet 1/0/3.
[DeviceB-GigabitEthernet1/0/3] mvrp enable
[DeviceB-GigabitEthernet1/0/3] quit
# Create VLAN 20.
[DeviceB] vlan 20
[DeviceB-vlan20] quit

```

Configuring Device C

```

# Enter MST region view.
<DeviceC> system-view
[DeviceC] stp region-configuration
# Configure the MST region name, VLAN-to-instance mappings, and revision level.
[DeviceC-mst-region] region-name example
[DeviceC-mst-region] instance 1 vlan 10
[DeviceC-mst-region] instance 2 vlan 20
[DeviceC-mst-region] revision-level 0
# Manually activate the MST region configuration.
[DeviceC-mst-region] active region-configuration
[DeviceC-mst-region] quit
# Globally enable the spanning tree feature.
[DeviceC] stp enable
# Globally enable MVRP.
[DeviceC] mvrp global enable
# Configure port GigabitEthernet 1/0/1 as a trunk port, and configure it to permit all VLANs.
[DeviceC] interface gigabitethernet 1/0/1
[DeviceC-GigabitEthernet1/0/1] port link-type trunk
[DeviceC-GigabitEthernet1/0/1] port trunk permit vlan all
# Enable MVRP on port GigabitEthernet 1/0/1.
[DeviceC-GigabitEthernet1/0/1] mvrp enable
[DeviceC-GigabitEthernet1/0/1] quit
# Configure port GigabitEthernet1/0/2 as a trunk port, and configure it to permit all VLANs.
[DeviceC] interface gigabitethernet 1/0/2
[DeviceC-GigabitEthernet1/0/2] port link-type trunk
[DeviceC-GigabitEthernet1/0/2] port trunk permit vlan all

```

```
# Enable MVRP on port GigabitEthernet1/0/2.
[DeviceC-GigabitEthernet1/0/2] mvrp enable
[DeviceC-GigabitEthernet1/0/2] quit
```

Configure Device D:

```
# Enter MST region view.
<DeviceD> system-view
[DeviceD] stp region-configuration

# Configure the MST region name, VLAN-to-instance mappings, and revision level.
[DeviceD-mst-region] region-name example
[DeviceD-mst-region] instance 1 vlan 10
[DeviceD-mst-region] instance 2 vlan 20
[DeviceD-mst-region] revision-level 0

# Manually activate the MST region configuration.
[DeviceD-mst-region] active region-configuration
[DeviceD-mst-region] quit

# Globally enable the spanning tree feature.
[DeviceD] stp enable

# Globally enable MVRP.
[DeviceD] mvrp global enable

# Configure port GigabitEthernet 1/0/1 as a trunk port, and configure it to permit VLANs 20 and 40.
[DeviceD] interface gigabitethernet 1/0/1
[DeviceD-GigabitEthernet1/0/1] port link-type trunk
[DeviceD-GigabitEthernet1/0/1] port trunk permit vlan 20 40

# Enable MVRP on port GigabitEthernet 1/0/1.
[DeviceD-GigabitEthernet1/0/1] mvrp enable
[DeviceD-GigabitEthernet1/0/1] quit

# Configure port GigabitEthernet1/0/2 as a trunk port, and configure it to permit VLAN 40.
[DeviceD] interface gigabitethernet 1/0/2
[DeviceD-GigabitEthernet1/0/2] port link-type trunk
[DeviceD-GigabitEthernet1/0/2] port trunk permit vlan 40

# Enable MVRP on port GigabitEthernet1/0/2.
[DeviceD-GigabitEthernet1/0/2] mvrp enable
[DeviceD-GigabitEthernet1/0/2] quit
```

Verifying the configuration

1. Verify the normal registration mode configuration:

Use the **display mvrp running-status** command to display the local MVRP VLAN information to verify whether the configuration takes effect.

Check the local VLAN information on Device A.

```
[DeviceA] display mvrp running-status
-----[MVRP Global Info]-----
Global Status      : Enabled
Compliance-GVRP   : False
```

```

----[GigabitEthernet1/0/1]----
Config Status           : Enabled
Running Status          : Enabled
Join Timer               : 20 (centiseconds)
Leave Timer               : 60 (centiseconds)
Periodic Timer           : 100 (centiseconds)
LeaveAll Timer           : 1000 (centiseconds)
Registration Type        : Normal
Local VLANs :
  1(default),

```

```

----[GigabitEthernet1/0/2]----
Config Status           : Enabled
Running Status          : Enabled
Join Timer               : 20 (centiseconds)
Leave Timer               : 60 (centiseconds)
Periodic Timer           : 100 (centiseconds)
LeaveAll Timer           : 1000 (centiseconds)
Registration Type        : Normal
Local VLANs :
  1(default),

```

```

----[GigabitEthernet1/0/3]----
Config Status           : Enabled
Running Status          : Enabled
Join Timer               : 20 (centiseconds)
Leave Timer               : 60 (centiseconds)
Periodic Timer           : 100 (centiseconds)
LeaveAll Timer           : 1000 (centiseconds)
Registration Type        : Normal
Local VLANs :
  1(default), 20,

```

The output shows that:

- Ports GigabitEthernet 1/0/1 and GigabitEthernet1/0/2 have learned only VLAN 1 through MVRP.
- Port GigabitEthernet 1/0/3 has learned VLAN 1 and dynamic VLAN 20 created on Device B through MVRP.

Check the local VLAN information on Device B.

```
[DeviceB] display mvrp running-status
```

```

-----[MVRP Global Info]-----
Global Status           : Enabled
Compliance-GVRP        : False

```

```

----[GigabitEthernet1/0/1]----
Config Status           : Enabled
Running Status          : Enabled
Join Timer               : 20 (centiseconds)
Leave Timer               : 60 (centiseconds)

```

```

Periodic Timer                : 100 (centiseconds)
LeaveAll Timer                 : 1000 (centiseconds)
Registration Type             : Normal
Local VLANs :
  1(default),

```

----[GigabitEthernet1/0/2]----

```

Config Status                 : Enabled
Running Status               : Enabled
Join Timer                   : 20 (centiseconds)
Leave Timer                   : 60 (centiseconds)
Periodic Timer               : 100 (centiseconds)
LeaveAll Timer                : 1000 (centiseconds)
Registration Type            : Normal
Local VLANs :
  1(default),

```

----[GigabitEthernet1/0/3]----

```

Config Status                 : Enabled
Running Status               : Enabled
Join Timer                   : 20 (centiseconds)
Leave Timer                   : 60 (centiseconds)
Periodic Timer               : 100 (centiseconds)
LeaveAll Timer                : 1000 (centiseconds)
Registration Type            : Normal
Local VLANs :
  1(default), 10,

```

The output shows that:

- Ports GigabitEthernet 1/0/1 and GigabitEthernet1/0/2 have learned only VLAN 1 through MVRP.
- Port GigabitEthernet 1/0/3 has learned VLAN 1 and dynamic VLAN 10 created on Device A through MVRP.

Check the local VLAN information on Device C.

```
[DeviceC] display mvrp running-status
```

-----[MVRP Global Info]-----

```

Global Status      : Enabled
Compliance-GVRP   : False

```

----[GigabitEthernet1/0/1]----

```

Config Status                 : Enabled
Running Status               : Enabled
Join Timer                   : 20 (centiseconds)
Leave Timer                   : 60 (centiseconds)
Periodic Timer               : 100 (centiseconds)
LeaveAll Timer                : 1000 (centiseconds)
Registration Type            : Normal
Local VLANs :
  1(default), 10,

```

```
----[GigabitEthernet1/0/2]----
```

```
Config Status           : Enabled
Running Status          : Enabled
Join Timer               : 20 (centiseconds)
Leave Timer              : 60 (centiseconds)
Periodic Timer          : 100 (centiseconds)
LeaveAll Timer           : 1000 (centiseconds)
Registration Type       : Normal
Local VLANs :
  1(default), 20,
```

The output shows that:

- Port GigabitEthernet 1/0/1 has learned VLAN 1 and dynamic VLAN 10 created on Device A through MVRP.
- Port GigabitEthernet1/0/2 has learned VLAN 1 and dynamic VLAN 20 created on Device B through MVRP.

Check the local VLAN information on Device D.

```
[DeviceD] display mvrp running-status
```

```
-----[MVRP Global Info]-----
```

```
Global Status      : Enabled
Compliance-GVRP   : False
```

```
----[GigabitEthernet1/0/1]----
```

```
Config Status           : Enabled
Running Status          : Enabled
Join Timer               : 20 (centiseconds)
Leave Timer              : 60 (centiseconds)
Periodic Timer          : 100 (centiseconds)
LeaveAll Timer           : 1000 (centiseconds)
Registration Type       : Normal
Local VLANs :
  1(default), 20,
```

```
----[GigabitEthernet1/0/2]----
```

```
Config Status           : Enabled
Running Status          : Enabled
Join Timer               : 20 (centiseconds)
Leave Timer              : 60 (centiseconds)
Periodic Timer          : 100 (centiseconds)
LeaveAll Timer           : 1000 (centiseconds)
Registration Type       : Normal
Local VLANs :
  1(default),
```

The output shows that:

- Port GigabitEthernet 1/0/1 has learned VLAN 1 and dynamic VLAN 20 created on Device B through MVRP.
- Port GigabitEthernet1/0/2 has learned only VLAN 1 through MVRP.

2. Change the registration mode and verify the configuration:

Set the MVRP registration mode to fixed on GigabitEthernet 1/0/3 of Device B, so that the dynamic VLANs that Device B learns in VLAN 1 are not de-registered.

Set the MVRP registration mode to fixed on GigabitEthernet 1/0/3.

```
[DeviceB] interface gigabitethernet 1/0/3
[DeviceB-GigabitEthernet1/0/3] mvrp registration fixed
[DeviceB-GigabitEthernet1/0/3] quit
```

Display the local MVRP VLAN information on GigabitEthernet 1/0/3.

```
[DeviceB] display mvrp running-status interface gigabitethernet 1/0/3
-----[MVRP Global Info]-----
Global Status      : Enabled
Compliance-GVRP   : False

----[GigabitEthernet1/0/3]----
Config Status      : Enabled
Running Status     : Enabled
Join Timer         : 20 (centiseconds)
Leave Timer        : 60 (centiseconds)
Periodic Timer     : 100 (centiseconds)
LeaveAll Timer     : 1000 (centiseconds)
Registration Type  : Fixed
Local VLANs :
  1(default), 10,
```

The output shows that the VLAN information on GigabitEthernet 1/0/3 is not changed after you set the MVRP registration mode to fixed on GigabitEthernet 1/0/3.

Delete VLAN 10 on Device A.

```
[DeviceA] undo vlan 10
```

Display the local MVRP VLAN information on GigabitEthernet 1/0/3.

```
[DeviceB] display mvrp running-status interface gigabitethernet 1/0/3
-----[MVRP Global Info]-----
Global Status      : Enabled
Compliance-GVRP   : False

----[GigabitEthernet1/0/3]----
Config Status      : Enabled
Running Status     : Enabled
Join Timer         : 20 (centiseconds)
Leave Timer        : 60 (centiseconds)
Periodic Timer     : 100 (centiseconds)
LeaveAll Timer     : 1000 (centiseconds)
Registration Type  : Fixed
Local VLANs :
  1(default), 10,
```

The output shows that the dynamic VLAN information on GigabitEthernet 1/0/3 is not changed after you set the MVRP registration mode to fixed on GigabitEthernet 1/0/3.

Index

[A](#) [B](#) [C](#) [D](#) [E](#) [G](#) [I](#) [L](#) [M](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [V](#)

A

Assigning a port to the isolation group, [58](#)

B

BPDU tunneling configuration examples, [113](#)

C

Configuration example for MVRP in normal registration mode, [252](#)

Configuration guidelines, [21](#)

Configuration prerequisites, [248](#)

Configuration prerequisites, [156](#)

Configuration procedure, [21](#)

Configuration procedure, [142](#)

Configuration procedure, [148](#)

Configuration procedure, [242](#)

Configuration restrictions and guidelines, [39](#)

Configuration restrictions and guidelines, [148](#)

Configuration restrictions and guidelines, [242](#)

Configuring a Layer 2 Ethernet interface, [7](#)

Configuring a loopback interface, [18](#)

Configuring a null interface, [19](#)

Configuring a port to operate in automatic voice VLAN assignment mode, [157](#)

Configuring a port to operate in manual voice VLAN assignment mode, [158](#)

Configuring an aggregate interface, [43](#)

Configuring an aggregation group, [39](#)

Configuring an MST region, [78](#)

Configuring basic QinQ, [180](#)

Configuring basic settings of a VLAN interface, [119](#)

Configuring basic VLAN settings, [118](#)

Configuring CDP compatibility, [231](#)

Configuring destination multicast MAC address for BPDUs, [112](#)

Configuring Digest Snooping, [92](#)

Configuring edge ports, [84](#)

Configuring GVRP functions, [168](#)

Configuring IP subnet-based VLANs, [138](#)

Configuring LLDP to advertise a specific voice VLAN, [233](#)

Configuring LLDP trapping, [235](#)

Configuring load sharing for link aggregation groups, [47](#)

Configuring MAC Information, [30](#)

Configuring MAC-based VLANs, [127](#)

Configuring many-to-one VLAN mapping, [206](#)

Configuring MRP timers, [250](#)

Configuring No Agreement Check, [94](#)

Configuring one-to-one VLAN mapping, [204](#)

Configuring path costs of ports, [85](#)

Configuring port-based VLANs, [121](#)

Configuring protection functions, [98](#)

Configuring protocol-based VLANs, [134](#)

Configuring QoS priority settings for voice traffic on an interface, [157](#)

Configuring selective QinQ, [181](#)

Configuring spanning tree timers, [82](#)

Configuring static, dynamic, and blackhole MAC address table entries, [23](#)

Configuring TC snooping, [96](#)

Configuring the aging timer for dynamic MAC address entries, [25](#)

Configuring the device priority, [80](#)

Configuring the GARP timers, [169](#)

Configuring the maximum hops of an MST region, [81](#)

Configuring the maximum port rate, [84](#)

Configuring the mode a port uses to recognize/send MSTP packets, [89](#)

Configuring the MVRP registration mode, [249](#)

Configuring the network diameter of a switched network, [81](#)

Configuring the port link type, [88](#)

Configuring the port priority, [88](#)

Configuring the root bridge or a secondary root bridge, [79](#)

Configuring the timeout factor, [83](#)

Configuring the TPID value in VLAN tags, [185](#)

Configuring two-to-two VLAN mapping, [208](#)

D

- Disabling MAC address learning, [24](#)
- Displaying and maintaining an Ethernet interface, [17](#)
- Displaying and maintaining Ethernet link aggregation, [50](#)
- Displaying and maintaining GVRP, [170](#)
- Displaying and maintaining isolate-user-VLAN, [149](#)
- Displaying and maintaining LLDP, [235](#)
- Displaying and maintaining loopback and null interfaces, [19](#)
- Displaying and maintaining MAC address tables, [28](#)
- Displaying and maintaining MVRP, [251](#)
- Displaying and maintaining service loopback groups, [243](#)
- Displaying and maintaining super VLAN, [144](#)
- Displaying and maintaining the isolation group, [58](#)
- Displaying and maintaining the spanning tree, [101](#)
- Displaying and maintaining VLAN, [141](#)
- Displaying and maintaining voice VLAN, [159](#)
- Dynamically advertising server-assigned VLANs through LLDP, [235](#)

E

- Enabling BPDU tunneling, [111](#)
- Enabling GVRP compatibility, [251](#)
- Enabling link-aggregation traffic redirection, [49](#)
- Enabling LLDP to automatically discover IP phones, [232](#)
- Enabling MVRP, [248](#)
- Enabling outputting port state transition information, [90](#)
- Enabling the spanning tree feature, [90](#)
- Ethernet interface naming conventions, [1](#)
- Ethernet link aggregation configuration examples, [50](#)
- Ethernet link aggregation configuration task list, [39](#)

G

- GVRP configuration examples, [170](#)
- GVRP configuration task list, [167](#)

I

- Isolate-user-VLAN configuration example, [150](#)

L

- LLDP configuration examples, [236](#)
- LLDP configuration task list, [226](#)

M

- MAC address table configuration example, [28](#)
- MAC Information configuration example, [32](#)
- MSTP, [68](#)
- MVRP configuration task list, [248](#)

O

- Overview(Configuring an isolate-user-VLAN), [147](#)
- Overview(Configuring MAC Information), [30](#)
- Overview(Configuring a service loopback group (available only on the 5500 EI)), [241](#)
- Overview(Configuring GVRP), [164](#)
- Overview(Configuring LLDP), [221](#)
- Overview(Configuring MVRP), [245](#)
- Overview(Configuring a voice VLAN), [153](#)
- Overview(Configuring the MAC address table), [22](#)
- Overview(Configuring QinQ), [176](#)
- Overview(Configuring BPDU tunneling), [109](#)
- Overview(Configuring VLANs), [116](#)
- Overview(Configuring Ethernet link aggregation), [33](#)
- Overview(Configuring VLAN mapping (available only on the 5500 EI)), [199](#)
- Overview(Configuring a super VLAN (available only on the 5500 EI)), [142](#)

P

- Performing basic LLDP configuration, [227](#)
- Performing general configurations, [1](#)
- Performing mCheck, [91](#)
- Port isolation configuration example, [59](#)
- Protocols and standards, [73](#)
- PVST, [67](#)

Q

- QinQ configuration examples, [186](#)
- QinQ configuration task list, [179](#)

R

- RSTP, [67](#)

S

- Service loopback group configuration example, [243](#)
- Setting the MTU for a Layer 3 Ethernet interface (available only on the 5500 EI), [16](#)
- Setting the spanning tree mode, [77](#)
- Spanning tree configuration examples, [102](#)
- Spanning tree configuration task list, [73](#)

STP, [60](#)

Super VLAN configuration example, [144](#)

V

VLAN mapping configuration examples, [212](#)

VLAN mapping configuration tasks, [203](#)

Voice VLAN configuration examples, [159](#)

Contents

Configuring ARP	1
Overview	1
ARP message format	1
ARP operation	2
ARP table	3
Configuring a static ARP entry	3
Configuring the maximum number of dynamic ARP entries for an interface	4
Setting the aging timer for dynamic ARP entries	4
Enabling dynamic ARP entry check	5
Configuring ARP quick update	5
Configuring multicast ARP	6
Displaying and maintaining ARP	6
ARP configuration examples	7
Static ARP entry configuration example	7
Multicast ARP configuration example	8
Configuring gratuitous ARP	11
Overview	11
Gratuitous ARP packet learning	11
Periodic sending of gratuitous ARP packets	11
Configuration guidelines	12
Configuration procedure	12
Configuring proxy ARP	13
Overview	13
Common proxy ARP	13
Local proxy ARP	13
Enabling common proxy ARP	14
Enabling local proxy ARP	14
Displaying and maintaining proxy ARP	15
Proxy ARP configuration examples	15
Common proxy ARP configuration example	15
Local proxy ARP configuration example in case of port isolation	16
Local proxy ARP configuration example in super VLAN(only available on the HP 5500 EI)	17
Local proxy ARP configuration example in isolate-user-VLAN	18
Configuring ARP snooping	21
Overview	21
Configuration procedure	21
Displaying and maintaining ARP snooping	21
Configuring IP addressing	22
Overview	22
IP address classes	22
Special IP addresses	23
Subnetting and masking	23
Assigning an IP address to an interface	24
Configuration guidelines	24
Configuration procedure	24
Configuration example	24
Configuring IP unnumbered(only available on the HP 5500 EI)	26

Overview	26
Configuration guidelines	26
Configuration prerequisites	26
Configuration procedure	27
Displaying and maintaining IP addressing	27
DHCP overview	28
DHCP address allocation	28
Dynamic IP address allocation process	29
IP address lease extension	29
DHCP message format	30
DHCP options	31
Common DHCP options	31
Custom options	31
Protocols and standards	35
Configuring DHCP server	36
Overview	36
DHCP address pool	36
IP address allocation sequence	37
DHCP server configuration task list	38
Configuring an address pool for the DHCP server	38
Configuration task list	38
Creating a DHCP address pool	39
Configuring address allocation mode for a common address pool	39
Configuring dynamic address allocation for an extended address pool	41
Configuring a domain name suffix for the client	42
Configuring DNS servers for the client	42
Configuring WINS servers and NetBIOS node type for the client	42
Configuring BIMS server information for the client	43
Configuring gateways for the client	43
Configuring Option 184 parameters for the client with voice service	44
Configuring the TFTP server and bootfile name for the client	44
Specifying a server's IP address for the DHCP client	45
Configuring self-defined DHCP options	45
Enabling DHCP	46
Enabling the DHCP server on an interface	46
Configuration guidelines	46
Configuration procedure	47
Applying an extended address pool on an interface	47
Configuring the DHCP server security functions	48
Configuration prerequisites	48
Enabling unauthorized DHCP server detection	48
Configuring IP address conflict detection	48
Enabling client offline detection	49
Enabling handling of Option 82	49
Configuration prerequisites	49
Enabling Option 82 handling	49
Specifying the threshold for sending trap messages	50
Configuration prerequisites	50
Configuration procedure	50
Setting the DSCP value for DHCP packets	50
Displaying and maintaining the DHCP server	50
DHCP server configuration examples	51
Static IP address assignment configuration example	51

Dynamic IP address assignment configuration example.....	53
Self-defined option configuration example.....	54
Troubleshooting DHCP server configuration.....	55
Symptom.....	55
Analysis.....	55
Solution.....	55
Configuring DHCP relay agent.....	56
Overview.....	56
Fundamentals.....	56
DHCP relay agent support for Option 82.....	57
DHCP relay agent configuration task list.....	58
Enabling DHCP.....	58
Enabling the DHCP relay agent on an interface.....	58
Correlating a DHCP server group with a relay agent interface.....	59
Configuration guidelines.....	59
Configuration procedure.....	59
Configuring the DHCP relay agent security functions.....	60
Configuring address check.....	60
Configuring periodic refresh of dynamic client entries.....	60
Enabling unauthorized DHCP server detection.....	61
Enabling DHCP starvation attack protection.....	61
Enabling offline detection.....	62
Configuring the DHCP relay agent to release an IP address.....	62
Configuring the DHCP relay agent to support Option 82.....	63
Configuration prerequisites.....	63
Configuration guidelines.....	63
Configuration procedure.....	63
Setting the DSCP value for DHCP packets.....	64
Displaying and maintaining the DHCP relay agent.....	64
DHCP relay agent configuration examples.....	65
DHCP relay agent configuration example.....	65
DHCP relay agent Option 82 support configuration example.....	66
Troubleshooting DHCP relay agent configuration.....	67
Symptom.....	67
Analysis.....	67
Solution.....	67
Configuring DHCP client.....	68
Configuration restrictions.....	68
Enabling the DHCP client on an interface.....	68
Setting the DSCP value for DHCP packets.....	68
Displaying and maintaining the DHCP client.....	69
DHCP client configuration example.....	69
Network requirements.....	69
Configuration procedure.....	69
Verifying the configuration.....	70
Configuring DHCP snooping.....	71
DHCP snooping functions.....	71
Ensuring that DHCP clients obtain IP addresses from authorized DHCP servers.....	71
Recording IP-to-MAC mappings of DHCP clients.....	71
Application environment of trusted ports.....	72
Configuring a trusted port connected to a DHCP server.....	72
Configuring trusted ports in a cascaded network.....	72
DHCP snooping support for Option 82.....	73

DHCP snooping configuration task list	74
Configuring DHCP snooping basic functions	75
Configuration guidelines	75
Configuration procedure	75
Configuring DHCP snooping to support Option 82	76
Configuration guidelines	76
Configuration procedure	76
Configuring DHCP snooping entries backup	77
Enabling DHCP starvation attack protection	78
Enabling DHCP-REQUEST message attack protection	79
Configuring DHCP packet rate limit	79
Configuration guidelines	79
Configuration procedure	79
Displaying and maintaining DHCP snooping	80
DHCP snooping configuration examples	80
DHCP snooping configuration example	80
DHCP snooping Option 82 support configuration example	81
Configuring BOOTP client	83
Overview	83
BOOTP application	83
Obtaining an IP address dynamically	83
Protocols and standards	83
Configuration restrictions	83
Configuring an interface to dynamically obtain an IP address through BOOTP	84
Displaying and maintaining BOOTP client configuration	84
BOOTP client configuration example	84
Network requirements	84
Configuration procedure	84
Configuring IPv4 DNS	85
Overview	85
Static domain name resolution	85
Dynamic domain name resolution	85
DNS proxy	86
DNS spoofing	87
Configuring the IPv4 DNS client	88
Configuring static domain name resolution	88
Configuring dynamic domain name resolution	88
Configuring the DNS proxy	89
Configuring DNS spoofing	90
Setting the DSCP value for DNS packets	90
Specifying the source interface for DNS packets	90
Displaying and maintaining IPv4 DNS	91
Static domain name resolution configuration example	91
Network requirements	91
Configuration procedure	91
Dynamic domain name resolution configuration example	92
Network requirements	92
Configuration procedure	93
Verifying the configuration	95
DNS proxy configuration example	95
Network requirements	95
Configuration procedure	96
Verifying the configuration	96

Troubleshooting IPv4 DNS configuration	97
Symptom	97
Solution	97
Configuring IRDP	98
Overview	98
Background	98
Working mechanism	98
Concepts	99
Protocols and standards	99
Configuration procedure	99
IRDP configuration example	100
Network requirements	100
Configuration procedure	101
Verifying the configuration	102
Configuring IP performance optimization	103
Enabling receiving and forwarding of directed broadcasts to a directly connected network	103
Enabling receiving of directed broadcasts to a directly connected network	103
Enabling forwarding of directed broadcasts to a directly connected network	103
Configuration example	104
Configuring TCP attributes	104
Configuring TCP path MTU discovery	104
Configuring the TCP send/receive buffer size	105
Configuring TCP timers	105
Configuring ICMP to send error packets	106
Advantages of sending ICMP error packets	106
Disadvantages of sending ICMP error packets	107
Configuration procedure	107
Displaying and maintaining IP performance optimization	108
Configuring UDP helper	109
Overview	109
Configuration restrictions and guidelines	109
Configuration procedure	109
Displaying and maintaining UDP helper	110
UDP helper configuration example	110
Network requirements	110
Configuration procedure	110
Configuring IPv6 basics	112
Overview	112
IPv6 features	112
IPv6 addresses	113
IPv6 neighbor discovery protocol	116
IPv6 path MTU discovery	118
IPv6 transition technologies	119
Protocols and standards	119
IPv6 basics configuration task list	120
Configuring basic IPv6 functions	121
Enabling IPv6	121
Configuring an IPv6 global unicast address	121
Configuring an IPv6 link-local address	123
Configure an IPv6 anycast address	124
Configuring IPv6 ND	125
Configuring a static neighbor entry	125

Configuring the maximum number of neighbors dynamically learned	125
Setting the age timer for ND entries in stale state	126
Configuring parameters related to RA messages	126
Configuring the maximum number of attempts to send an NS message for DAD	128
Configuring ND snooping	129
Enabling ND proxy	131
Configuring path MTU discovery	133
Configuring a static path MTU for a specified IPv6 address	133
Configuring the aging time for dynamic path MTUs	133
Configuring IPv6 TCP properties	133
Configuring ICMPv6 packet sending	134
Configuring the maximum ICMPv6 error packets sent in an interval	134
Enabling replying to multicast echo requests	134
Enabling sending of ICMPv6 time exceeded messages	135
Enabling sending of ICMPv6 destination unreachable messages	135
Displaying and maintaining IPv6 basics configuration	136
IPv6 basics configuration example	137
Network requirements	137
Configuration procedure	138
Verifying the configuration	139
Troubleshooting IPv6 basics configuration	142
Symptom	142
Solution	142
DHCPv6 overview	143
Introduction to DHCPv6	143
DHCPv6 address/prefix assignment	143
Rapid assignment involving two messages	143
Assignment involving four messages	143
Address/prefix lease renewal	144
Configuring stateless DHCPv6	145
Operation	145
Protocols and standards	146
Configuring DHCPv6 server	147
Overview	147
Concepts	147
Prefix selection process	148
DHCPv6 server configuration task list	148
Enabling the DHCPv6 server	149
Creating a prefix pool	149
Configuring a DHCPv6 address pool	149
Configuration restrictions and guidelines	149
Configuration procedure	149
Applying the address pool to an interface	150
Setting the DSCP value for DHCPv6 packets	151
Displaying and maintaining the DHCPv6 server	151
DHCPv6 server configuration example	151
Network requirements	151
Configuration considerations	152
Configuration procedure	152
Verifying the configuration	153
Configuring DHCPv6 relay agent	155
Overview	155
DHCPv6 relay agent operation	155

Configuring the DHCPv6 relay agent	156
Configuration guidelines	156
Configuration procedure	156
Setting the DSCP value for DHCPv6 packets	157
Displaying and maintaining the DHCPv6 relay agent	157
DHCPv6 relay agent configuration example	157
Network requirements	157
Configuration procedure	158
Verifying the configuration	158
Configuring DHCPv6 client	160
Overview	160
Configuring the DHCPv6 client	160
Configuration prerequisites	160
Configuration guidelines	160
Configuration procedure	160
Setting the DSCP value for DHCPv6 packets	160
Displaying and maintaining the DHCPv6 client	161
Stateless DHCPv6 configuration example	161
Network requirements	161
Configuration procedure	161
Verifying the configuration	162
Configuring DHCPv6 snooping	164
Overview	164
Ensuring that DHCPv6 clients obtain IPv6 addresses from authorized DHCPv6 servers	164
Recording IP-to-MAC mappings of DHCPv6 clients	165
Enabling DHCPv6 snooping	165
Configuring a DHCPv6 snooping trusted port	165
Configuring the maximum number of DHCPv6 snooping entries an interface can learn	166
Displaying and maintaining DHCPv6 snooping	166
DHCPv6 snooping configuration example	166
Network requirements	166
Configuration procedure	167
Verifying the configuration	167
Configuring IPv6 DNS	168
Overview	168
Configuring the IPv6 DNS client	168
Configuring static domain name resolution	168
Configuring dynamic domain name resolution	168
Setting the DSCP value for IPv6 DNS packets	169
Displaying and maintaining IPv6 DNS	169
Static domain name resolution configuration example	170
Network requirements	170
Configuration procedure	170
Dynamic domain name resolution configuration example	171
Network requirements	171
Configuration procedure	171
Verifying the configuration	174
Configuring tunneling(only available on the HP 5500 EI).....	176
Overview	176
IPv6 over IPv4 tunneling	176
Protocols and standards	178
Tunneling configuration task list	178

Configuring a tunnel interface	179
Configuration guidelines	179
Configuration procedure	179
Configuring an IPv6 manual tunnel	180
Configuration prerequisites	180
Configuration guidelines	180
Configuration procedure	180
Configuration example	181
Configuring a 6to4 tunnel	185
Configuration prerequisites	185
Configuration guidelines	185
Configuration procedure	185
Configuration example	186
Configuring an ISATAP tunnel	188
Configuration prerequisites	188
Configuration guidelines	189
Configuration procedure	189
Configuration example	190
Displaying and maintaining tunneling configuration	192
Troubleshooting tunneling configuration	193
Symptom	193
Solution	193
Index	194

Configuring ARP

Only HP 5500 EI switches support Layer 3 Ethernet port configuration.

You can use the **port link-mode** command to set an Ethernet port to operate in bridge (Layer 2) or route mode (Layer 3) (see *Layer 2—LAN Switching Configuration Guide*).

Overview

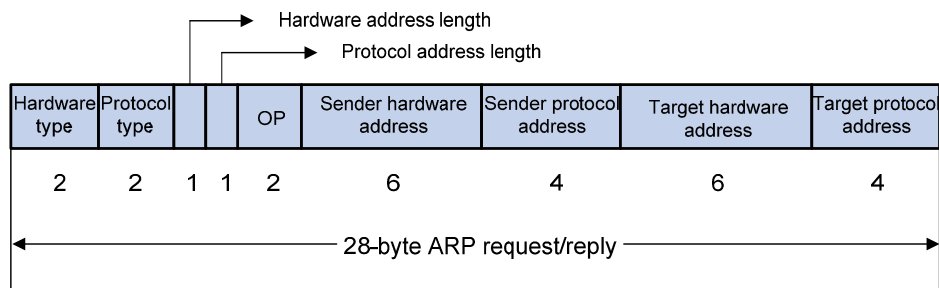
The Address Resolution Protocol (ARP) is used to resolve an IP address into a physical address (Ethernet MAC address, for example).

In an Ethernet LAN, a device uses ARP to resolve the IP address of the next hop to the corresponding MAC address.

ARP message format

ARP messages include ARP requests and ARP replies. Figure 1 shows the format of the ARP request/reply. Numbers in the figure refer to field lengths.

Figure 1 ARP message format



ARP message fields:

- **Hardware type**—The hardware address type. Value 1 represents Ethernet.
- **Protocol type**—The type of the protocol address to be mapped. The hexadecimal value 0x0800 represents IP.
- **Hardware address length and protocol address length**—Length, in bytes, of a hardware address and a protocol address. For an Ethernet address, the value of the hardware address length field is 6. For an IPv4 address, the value of the protocol address length field is 4.
- **OP**—Operation code, which describes type of the ARP message. Value 1 represents an ARP request, and value 2 represents an ARP reply.
- **Sender hardware address**—Hardware address of the device sending the message.
- **Sender protocol address**—Protocol address of the device sending the message.
- **Target hardware address**—Hardware address of the device to which the message is being sent.
- **Target protocol address**—Protocol address of the device to which the message is being sent.

ARP operation

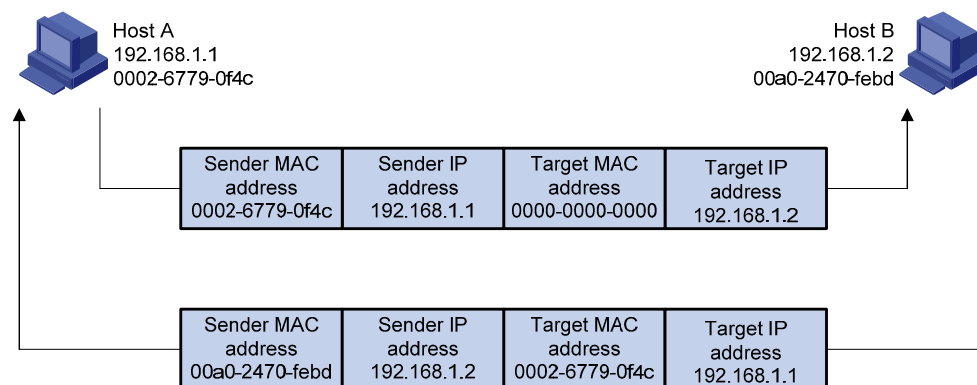
If Host A and Host B are on the same subnet and Host A sends a packet to Host B, as shown in [Figure 2](#), the resolution process is:

1. Host A looks in its ARP table to see whether there is an ARP entry for Host B. If yes, Host A uses the MAC address in the entry to encapsulate the IP packet into a data link layer frame and sends the frame to Host B.
2. If Host A finds no entry for Host B, Host A buffers the packet and broadcasts an ARP request using the following information:
 - **Source IP address and source MAC address**—Host A's own IP address and the MAC address
 - **Target IP address**—Host B's IP address
 - **Target MAC address**—An all-zero MAC address

All hosts on this subnet can receive the broadcast request, but only the requested host (Host B) processes the request.

3. Host B compares its own IP address with the target IP address in the ARP request. If they are the same, Host B:
 - a. Adds the sender IP address and sender MAC address into its ARP table.
 - b. Encapsulates its MAC address into an ARP reply.
 - c. Unicasts the ARP reply to Host A.
4. After receiving the ARP reply, Host A:
 - a. Adds the MAC address of Host B to its ARP table.
 - b. Encapsulates the MAC address into the packet and sends it to Host B.

Figure 2 ARP address resolution process



If Host A and Host B are on different subnets, the resolution process is as follows:

1. Host A sends an ARP request to the gateway. The target IP address in the ARP request is the IP address of the gateway.
2. After obtaining the MAC address of the gateway from an ARP reply, Host A sends the packet to the gateway.
3. If the gateway maintains the ARP entry of Host B, it forwards the packet to Host B directly; if not, it broadcasts an ARP request, in which the target IP address is the IP address of Host B.
4. After obtaining the MAC address of Host B, the gateway sends the packet to Host B.

ARP table

An ARP table stores dynamic and static ARP entries.

Dynamic ARP entry

ARP automatically creates and updates dynamic entries. A dynamic ARP entry is removed when its aging timer expires or the output interface goes down, and it can be overwritten by a static ARP entry.

Static ARP entry

A static ARP entry is manually configured and maintained. It does not age out, and cannot be overwritten by a dynamic ARP entry.

Static ARP entries protect communication between devices, because attack packets cannot modify the IP-to-MAC mapping in a static ARP entry.

Static ARP entries can be classified into long and short ARP entries.

- To configure a long static ARP entry, specify the IP address, MAC address, VLAN, and output interface. A long static ARP entry is directly used for forwarding matching packets. To allow communication with a host using a fixed IP-to-MAC mapping through a specific interface in a specific VLAN, configure a long static ARP entry for it.
- To configure a short static ARP entry, you only need to specify the IP address and MAC address.
 - If the output interface is a Layer 3 Ethernet port, the short ARP entry can be directly used for forwarding matching packets.
 - If the output interface is a VLAN interface, the device first sends an ARP request whose target IP address is the IP address of the short entry. If the sender IP and MAC addresses in the received ARP reply match the IP and MAC addresses of the short static ARP entry, the device adds the interface receiving the ARP reply to the short static ARP entry, and then uses the resolved entry to forward the matching IP packets.

To communicate with a host by using a fixed IP-to-MAC mapping, configure a short static ARP entry for it.

Configuring a static ARP entry

A static ARP entry is effective when the device it corresponds to works normally. However, when a VLAN or VLAN interface is deleted, any static ARP entry corresponding to it will also be deleted (if it is a long static ARP entry) or will become unresolved (if it is a short and resolved static ARP entry).

Follow these guidelines when you configure a long static ARP entry:

- The *vlan-id* argument must be the ID of an existing VLAN where the ARP entry resides. The specified Ethernet interface must belong to that VLAN. The VLAN interface of the VLAN must be created.
- The IP address of the VLAN interface of the VLAN specified by the *vlan-id* argument must belong to the same subnet as the IP address specified by the *ip-address* argument.

To configure a static ARP entry:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Configure a static ARP entry.	<ul style="list-style-type: none"> Configure a long static ARP entry: arp static <i>ip-address mac-address vlan-id interface-type interface-number</i> [vpn-instance <i>vpn-instance-name</i>] Configure a short static ARP entry: arp static <i>ip-address mac-address</i> [vpn-instance <i>vpn-instance-name</i>] 	<p>Use either command.</p> <p>Only HP 5500 EI Switch Series supports the vpn-instance <i>vpn-instance-name</i> option</p>

Configuring the maximum number of dynamic ARP entries for an interface

An interface can dynamically learn ARP entries, so it may hold too many ARP entries. To solve this problem, you can set the maximum number of dynamic ARP entries that an interface can learn. When the maximum number is reached, the interface stops learning ARP entries.

A Layer 2 interface can learn an ARP entry only when both its maximum number and the VLAN interface's maximum number are not reached.

To set the maximum number of dynamic ARP entries that an interface can learn:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type interface-number</i>	N/A
3. Set the maximum number of dynamic ARP entries that the interface can learn.	arp max-learning-number <i>number</i>	<p>Optional.</p> <p>By default, a Layer 2 interface does not limit the number of dynamic ARP entries. A Layer 3 interface can learn up to 2048 ARP entries for HP 5500 SI Switch Series, and up to 8192 ARP entries for HP 5500 EI Switch Series.</p> <p>If the value of the <i>number</i> argument is set to 0, the interface is disabled from learning dynamic ARP entries.</p>

Setting the aging timer for dynamic ARP entries

Each dynamic ARP entry in the ARP table has a limited lifetime, called aging timer. The aging timer of a dynamic ARP entry is reset each time the dynamic ARP entry is updated. Dynamic ARP entries that are not updated before their aging timers expire are deleted from the ARP table.

To set the age timer for dynamic ARP entries:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Set the age timer for dynamic ARP entries.	arp timer aging <i>aging-time</i>	Optional. 20 minutes by default.

Enabling dynamic ARP entry check

The dynamic ARP entry check function controls whether the device supports dynamic ARP entries with multicast MAC addresses.

When dynamic ARP entry check is enabled, the device cannot learn dynamic ARP entries containing multicast MAC addresses.

When dynamic ARP entry check is disabled, the device can learn dynamic ARP entries containing multicast MAC addresses.

To enable dynamic ARP entry check:

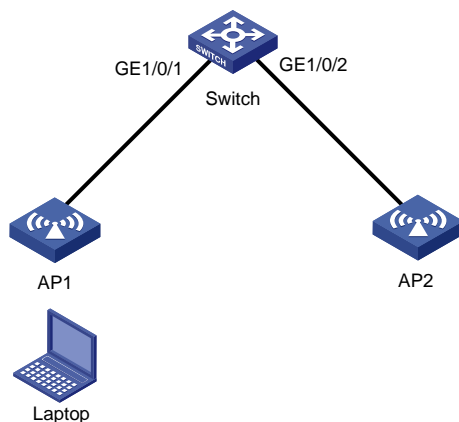
Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable dynamic ARP entry check.	arp check enable	Optional. Enabled by default.

Configuring ARP quick update

HP recommends enabling ARP quick update in WLANs only.

As shown in [Figure 3](#), the laptop frequently roams between AP 1 and AP 2. This affects the mapping between its MAC address and output interface on the switch. If the switch does not update its ARP table immediately after the output interface changes, it may fail to communicate with the laptop.

Figure 3 ARP quick update application scenario



With ARP quick update enabled, the switch updates the corresponding ARP entry immediately after the change of the mapping between a MAC address and an output interface to ensure nonstop data forwarding.

To enable ARP quick update:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable ARP quick update.	mac-address station-move quick-notify enable	Optional. Disabled by default.

Configuring multicast ARP

Microsoft Network Load Balancing (NLB) is a load balancing technology for server clustering developed on Windows Server.

NLB supports load sharing and redundancy among servers within a cluster. To implement fast failover, NLB requires that the switch forwards network traffic to all servers or specified servers in the cluster, and each server filters out unexpected traffic. In a medium or small data center that uses the Windows Server operating system, the proper cooperation of the switch and NLB is very important. For more information about NLB, see the related documents of Windows Sever.

Microsoft NLB provides the following packet sending modes to make the switch forward network traffic to all servers or specified servers:

- **Unicast mode**—NLB assigns each cluster member a common MAC address, which is the cluster MAC address, and changes the source MAC address of each sent packet. Thus, the switch cannot add the cluster MAC address to its MAC table. In addition, because the cluster MAC address is unknown to the switch, packets destined to it are forwarded on all the ports of the switch.
- **Multicast mode**—NLB uses a multicast MAC address that is a virtual MAC address for network communication, for example 0300-5e11-1111.
- **Internet Group Management Protocol (IGMP) multicast mode**—The switch sends packets only out of the ports that connect to the cluster members rather than all ports.

NOTE:

Multicast ARP is applicable to only multicast-mode NLB.

To configure multicast ARP:

Step	Command	Remarks
1. Disable the ARP entry check function.	undo arp check enable	N/A
2. Configure a static ARP entry.	arp static <i>ip-address mac-address vlan-id interface-type interface-number</i> [vpn-instance <i>vpn-instance-name</i>]	Optional. Only HP 5500 EI Switch Series supports the vpn-instance <i>vpn-instance-name</i> option.
3. Configure a static multicast MAC address entry.	mac-address multicast <i>mac-address interface interface-list vlan vlan-id</i>	See <i>IP Multicast Command Reference</i> .

Displaying and maintaining ARP

CAUTION:

- Clearing ARP entries from the ARP table may cause communication failures.
- The verbose keyword is available only on the HP 5500 EI Switch Series.

Task	Command	Remarks
Display ARP entries in the ARP table.	display arp [[all dynamic static] [slot <i>slot-number</i>] vlan <i>vlan-id</i> interface <i>interface-type interface-number</i>] [count verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the ARP entry for a specified IP address.	display arp <i>ip-address</i> [slot <i>slot-number</i>] [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the ARP entries for a specified VPN instance (only available on the HP 5500 EI).	display arp vpn-instance <i>vpn-instance-name</i> [count] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the age timer for dynamic ARP entries.	display arp timer aging [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear ARP entries from the ARP table.	reset arp { all dynamic static slot <i>slot-number</i> interface <i>interface-type interface-number</i> }	Available in user view

ARP configuration examples

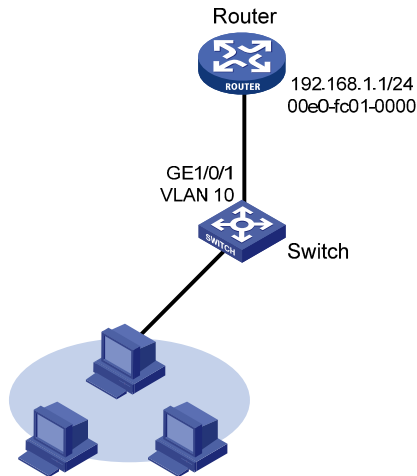
Static ARP entry configuration example

Network requirements

As shown in [Figure 4](#), hosts are connected to the switch, which is connected to the router through interface GigabitEthernet 1/0/1 in VLAN 10. The IP and MAC addresses of the router are 192.168.1.1/24 and 00e0-fc01-0000 respectively.

To prevent malicious users from attacking the switch and enhance security for communications between the router and switch, configure a static ARP entry for the router on the switch.

Figure 4 Network diagram



Configuration procedure

Configure the switch:

Create VLAN 10.

```
<Switch> system-view
[Switch] vlan 10
[Switch-vlan10] quit
```

Add interface GigabitEthernet 1/0/1 to VLAN 10.

```
[Switch] interface GigabitEthernet 1/0/1
[Switch-GigabitEthernet1/0/1] port link-type trunk
[Switch-GigabitEthernet1/0/1] port trunk permit vlan 10
[Switch-GigabitEthernet1/0/1] quit
```

Create interface VLAN-interface 10 and configure its IP address.

```
[Switch] interface vlan-interface 10
[Switch-vlan-interface10] ip address 192.168.1.2 24
[Switch-vlan-interface10] quit
```

Configure a static ARP entry that has IP address 192.168.1.1, MAC address 00e0-fc01-0000, and output interface GigabitEthernet 1/0/1 in VLAN 10.

```
[Switch] arp static 192.168.1.1 00e0-fc01-0000 10 GigabitEthernet 1/0/1
```

Display information about static ARP entries.

```
[Switch] display arp static
```

IP Address	MAC Address	VLAN ID	Interface	Aging Type
192.168.1.1	00e0-fc01-0000	10	GE1/0/1	N/A S

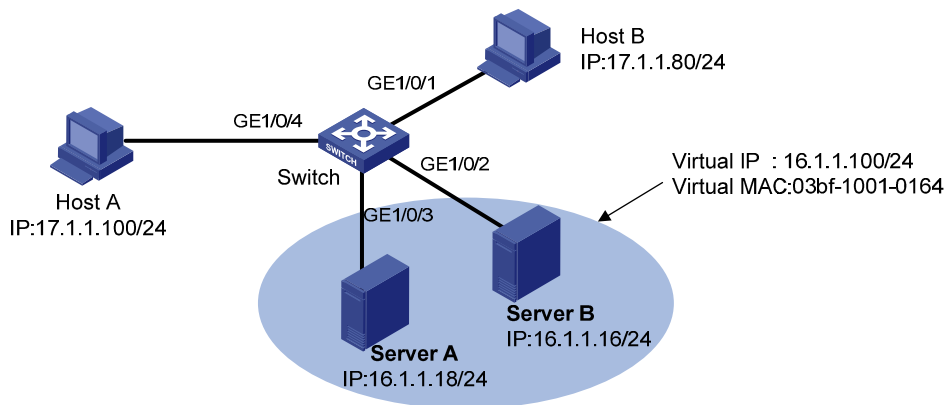
Multicast ARP configuration example

Network requirements

As shown in [Figure 5](#), a small data center uses Microsoft multicast-mode NLB. To enable the switches to cooperate with NLB, configure the following:

- Add GigabitEthernet 1/0/2 and GigabitEthernet 1/0/3 into VLAN 1, and specify IP address 16.1.1.30/24 for VLAN-interface 1.
- Add GigabitEthernet 1/0/1 and GigabitEthernet 1/0/4 into VLAN 2, and specify IP address 17.1.1.1/24 for VLAN-interface 2.
- Specify 17.1.1.1/24 as the default gateway of Host A and Host B.
- Specify 16.1.1.30/24 as the default gateway of Server A and Server B.
- Disable the ARP entry check function so that the switch can learn dynamic ARP entries containing multicast MAC addresses.
- Configure a static multicast MAC address entry so that only interfaces GigabitEthernet 1/0/2 and GigabitEthernet 1/0/3 can receive multicast information.

Figure 5 Network diagram



Configuration procedure

This example only describes multicast ARP configuration on the switch, and is only applicable to multicast NLB. For NLB configuration on the servers, see the related documents of the Windows Server.

Specify an IP address for VLAN-interface 2.

```
<Switch> system-view
[Switch] vlan 2
[Switch-vlan2] port GigabitEthernet 1/0/4
[Switch-vlan2] port GigabitEthernet 1/0/1
[Switch-vlan2] quit
[Switch] interface vlan-interface 2
[Switch-Vlan-interface2] ip address 17.1.1.1 255.255.255.0
[Switch-Vlan-interface2] quit
```

Specify an IP address for VLAN-interface 1.

```
[Switch] interface vlan-interface 1
[Switch-Vlan-interface1] ip address 16.1.1.30 255.255.255.0
[Switch-Vlan-interface1] quit
```

Disable the ARP entry check function.

```
[Switch] undo arp check enable
```

Configure a static multicast MAC address entry.

```
[Switch] mac-address multicast 03bf-1001-0164 interface GigabitEthernet 1/0/2 GigabitEthernet 1/0/3 vlan 1
```

Verifying the configuration

- **NLB load sharing**—Enables the FTP server function of Server A and Server B. Host A and Host B send requests to the virtual IP address and each of them logs in to a different server.
- **NLB redundancy**—Disables the network interface card of Server A. Host A and Host B send requests to the virtual IP address and both log in to the FTP server on Server B.

Configuring gratuitous ARP

Overview

In a gratuitous ARP packet, the sender IP address and the target IP address are the IP address of the sending device.

A device sends a gratuitous ARP packet for either of the following purposes:

- Determine whether its IP address is already used by another device. If the IP address is already used, the device is informed of the conflict by an ARP reply.
- Inform other devices of a change of its MAC address.

Gratuitous ARP packet learning

This feature enables a device to create or update ARP entries by using the sender IP and MAC addresses in received gratuitous ARP packets.

With this feature disabled, the device uses received gratuitous ARP packets to update existing ARP entries only.

Periodic sending of gratuitous ARP packets

Enabling a device to periodically send gratuitous ARP packets helps downstream devices update their corresponding ARP entries or MAC entries in time. This feature can be used to:

- Prevent gateway spoofing.

When an attacker sends forged gratuitous ARP packets to the hosts on a network, the traffic destined for the gateway from the hosts is sent to the attacker instead. As a result, the hosts cannot access the external network.

To prevent gateway spoofing attacks, enable the gateway to send gratuitous ARP packets containing its primary IP address and manually configured secondary IP addresses at a specific interval, so hosts can learn correct gateway address information.
- Prevent ARP entries from aging out.

If network traffic is heavy or if a host's CPU usage is high on a host, received ARP packets may be discarded or not be processed in time. Eventually, the dynamic ARP entries on the receiving host age out, and the traffic between the host and the corresponding devices is interrupted until the host re-creates the ARP entries.

To prevent this problem, enable the gateway to send gratuitous ARP packets periodically. The gratuitous ARP packets contain the gateway's primary IP address or one of its manually configured secondary IP addresses, so the receiving host can update ARP entries in time, ensuring traffic continuity.
- Prevent the virtual IP address of a VRRP group from being used by a host.

The master router of a VRRP group can periodically send gratuitous ARP packets to the hosts on the local network, so that the hosts can update local ARP entries and avoid using the virtual IP address of the VRRP group.

If the virtual IP address of the VRRP group is associated with a virtual MAC address, the sender MAC address in the gratuitous ARP packet takes the virtual MAC address of the virtual router. If the virtual IP address of the VRRP group is associated with the real MAC address of an interface, the sender MAC address in the gratuitous ARP packet takes the MAC address of the interface on the master router in the VRRP group.

For more information about VRRP, see *High Availability Configuration Guide*.

Configuration guidelines

Follow these guidelines when you configure gratuitous ARP:

- You can enable periodic sending of gratuitous ARP packets in VLAN interface view or Layer 3 Ethernet port view.
- You can enable periodic sending of gratuitous ARP packets on a maximum of 1024 interfaces.
- Periodic sending of gratuitous ARP packets takes effect only when the link of the enabled interface goes up and an IP address has been assigned to the interface.
- If you change the interval for sending gratuitous ARP packets, the configuration is effective at the next sending interval.
- The frequency of sending gratuitous ARP packets may be much lower than is expected if this function is enabled on multiple interfaces, if each interface is configured with multiple secondary IP addresses, or if a small sending interval is configured in such cases.

Configuration procedure

To configure gratuitous ARP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable learning of gratuitous ARP packets.	gratuitous-arp-learning enable	Optional. Enabled by default.
3. Enable the device to send gratuitous ARP packets upon receiving ARP requests from another subnet.	gratuitous-arp-sending enable	By default, a device does not send gratuitous ARP packets upon receiving ARP requests from another subnet.
4. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
5. Enable periodic sending of gratuitous ARP packets and set the sending interval.	arp send-gratuitous-arp [interval <i>milliseconds</i>]	Disabled by default.

Configuring proxy ARP

Overview

Proxy ARP enables a device on a network to answer ARP requests for an IP address not on that network. With proxy ARP, hosts on different broadcast domains can communicate with each other as they do on the same network.

Proxy ARP includes common proxy ARP and local proxy ARP.

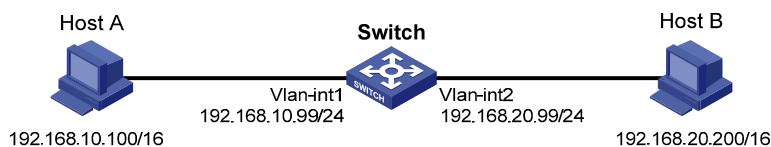
- **Common proxy ARP**—Allows communication between hosts that connect to different Layer-3 interfaces and reside in different broadcast domains.
- **Local proxy ARP**—Allows communication between hosts that connect to the same Layer-3 interface and reside in different broadcast domains.

Common proxy ARP

A proxy ARP enabled device allows hosts that reside on different subnets to communicate.

As shown in [Figure 6](#), Switch connects to two subnets through VLAN-interface 1 and VLAN-interface 2. The IP addresses of the two interfaces are 192.168.10.99/24 and 192.168.20.99/24. Host A and Host B are assigned the same prefix 192.168.0.0. Host A connects to VLAN-interface 1 and Host B connects to VLAN-interface 2.

Figure 6 Application environment of proxy ARP



Because Host A and Host B have the same prefix 192.168.0.0, Host A considers that Host B is on the same network, and it broadcasts an ARP request for the MAC address of Host B. However, Host B cannot receive this request because it is in a different broadcast domain.

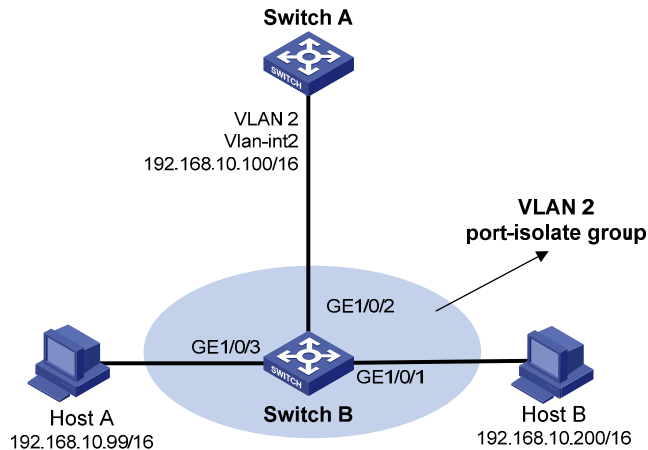
You can enable proxy ARP on VLAN-interface 1 of the switch so that the switch can reply to the ARP request from Host A with the MAC address of VLAN-interface 1, and forward packets sent from Host A to Host B. In this case, the switch acts as a proxy of Host B.

A main advantage of proxy ARP is that you can enable it on a single switch without disturbing routing tables of other routers in the network. Proxy ARP acts as the gateway for hosts that are not configured with a default gateway or do not have routing capability.

Local proxy ARP

As shown in [Figure 7](#), Host A and Host B belong to VLAN 2, but are isolated at Layer 2. Host A connects to GigabitEthernet 1/0/3 while Host B connects to GigabitEthernet 1/0/1. Enable local proxy ARP on Switch A to allow Layer 3 communication between the two hosts.

Figure 7 Application environment of local proxy ARP



Enable local proxy ARP in one of the following cases:

- Hosts connecting to different isolated Layer 2 ports in the same VLAN need to communicate at Layer 3.
- If a super VLAN is configured, hosts in different sub VLANs of the super VLAN need to communicate at Layer 3.
- If an isolate-user-VLAN is configured, hosts in different secondary VLANs of the isolate-user-VLAN need to communicate at Layer 3.

Enabling common proxy ARP

To enable common proxy ARP in VLAN interface view/Layer 3 Ethernet port view/Layer 3 aggregate interface view:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type interface-number</i>	N/A
3. Enable proxy ARP.	proxy-arp enable	Disabled by default

Enabling local proxy ARP

To enable local proxy ARP in VLAN interface view/Layer 3 Ethernet port view/Layer 3 aggregate interface view:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type interface-number</i>	N/A
3. Enable local proxy ARP.	local-proxy-arp enable [ip-range <i>startIP to endIP</i>]	Disabled by default

Displaying and maintaining proxy ARP

Task	Command	Remarks
Display whether proxy ARP is enabled.	display proxy-arp [interface <i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display whether local proxy ARP is enabled.	display local-proxy-arp [interface <i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Proxy ARP configuration examples

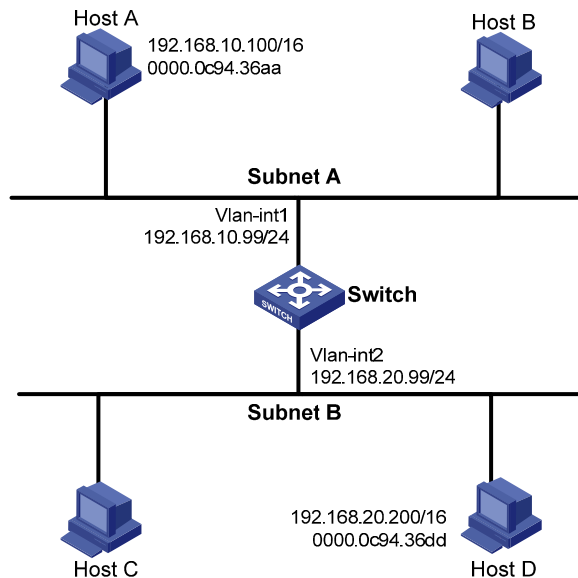
Common proxy ARP configuration example

Network requirements

As shown in [Figure 8](#), Host A and Host D have the same IP prefix and mask (IP addresses of Host A and Host D are 192.168.10.100/16 and 192.168.20.200/16 respectively), but they are located on different subnets separated by the switch (Host A belongs to VLAN 1 while Host D belongs to VLAN 2). As a result, Host D cannot receive or respond to any ARP request from Host A.

You must configure proxy ARP on the switch to enable communication between the two hosts.

Figure 8 Network diagram



Configuration procedure

```
# Create VLAN 2.  
<Switch> system-view  
[Switch] vlan 2  
[Switch-vlan2] quit
```

```

# Specify the IP address of interface VLAN-interface 1.
[Switch] interface vlan-interface 1
[Switch-Vlan-interface1] ip address 192.168.10.99 255.255.255.0

# Enable proxy ARP on interface VLAN-interface 1.
[Switch-Vlan-interface1] proxy-arp enable
[Switch-Vlan-interface1] quit

# Specify the IP address of interface VLAN-interface 2.
[Switch] interface vlan-interface 2
[Switch-Vlan-interface2] ip address 192.168.20.99 255.255.255.0

# Enable proxy ARP on interface VLAN-interface 2.
[Switch-Vlan-interface2] proxy-arp enable

```

After completing preceding configurations, use the **ping** command to verify the connectivity between Host A and Host D.

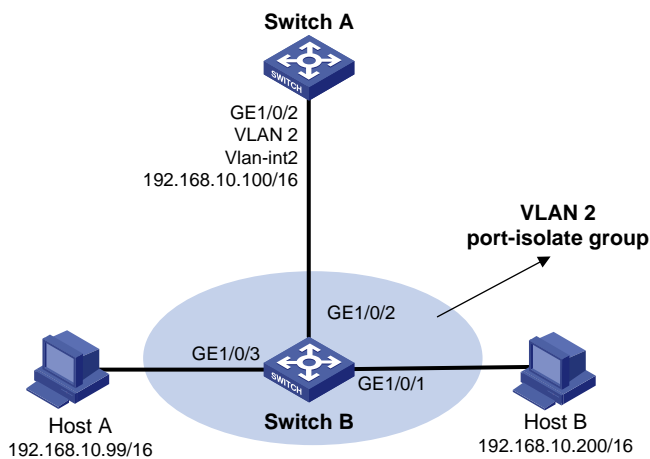
Local proxy ARP configuration example in case of port isolation

Network requirements

As shown in Figure 9, Host A and Host B belong to the same VLAN, and connect to Switch B via GigabitEthernet 1/0/3 and GigabitEthernet 1/0/1 respectively. Switch B connects to Switch A via GigabitEthernet 1/0/2.

Configure port isolation on GigabitEthernet 1/0/3 and GigabitEthernet 1/0/1 of Switch B to isolate Host A from Host B at Layer 2. Enable local proxy ARP on Switch A to allow communication between Host A and Host B at Layer 3.

Figure 9 Network diagram



Configuration procedure

1. Configure Switch B:
 - # Add GigabitEthernet 1/0/3, GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 to VLAN 2. Configure port isolation on Host A and Host B.

```

<SwitchB> system-view
[SwitchB] vlan 2
[SwitchB-vlan2] port GigabitEthernet 1/0/3

```

```

[SwitchB-vlan2] port GigabitEthernet 1/0/1
[SwitchB-vlan2] port GigabitEthernet 1/0/2
[SwitchB-vlan2] quit
[SwitchB] interface GigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] port-isolate enable
[SwitchB-GigabitEthernet1/0/3] quit
[SwitchB] interface GigabitEthernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] port-isolate enable
[SwitchB-GigabitEthernet1/0/1] quit

```

2. Configure Switch A:

Create VLAN 2, and add GigabitEthernet 1/0/2 to VLAN 2.

```

<SwitchA> system-view
[SwitchA] vlan 2
[SwitchA-vlan2] port GigabitEthernet 1/0/2
[SwitchA-vlan2] quit
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ip address 192.168.10.100 255.255.0.0

```

From Host A, ping Host B. The ping operation is unsuccessful because they are isolated at Layer 2.

Configure local proxy ARP to allow communication between Host A and Host B at Layer 3.

```

[SwitchA-Vlan-interface2] local-proxy-arp enable

```

From Host A, ping Host B. The ping operation is successful after the configuration.

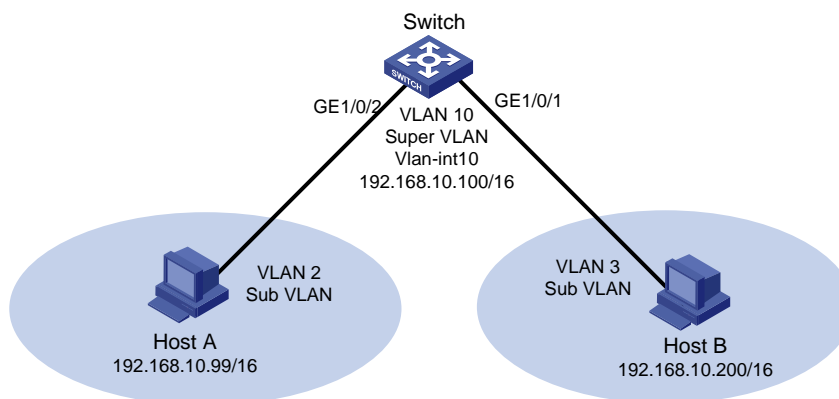
Local proxy ARP configuration example in super VLAN(only available on the HP 5500 EI)

Network requirements

Figure 10 shows a super VLAN, VLAN 10, with the interface IP address 192.168.10.100/16 and sub-VLANs (VLAN 2 and VLAN 3). GigabitEthernet 1/0/2 belongs to VLAN 2 and GigabitEthernet 1/0/1 belongs to VLAN 3. Host A belongs to VLAN 2 and connects to GigabitEthernet 1/0/2 of the switch. Host B belongs to VLAN 3 and connects to GigabitEthernet 1/0/1 of the switch.

As Host A and Host B belong to different Sub-VLANs, they are isolated at Layer 2. Configure local proxy ARP on the switch to allow Layer 3 communication between Host A and Host B.

Figure 10 Network diagram



Configuration procedure

Create the super VLAN and the sub-VLANs. Add GigabitEthernet 1/0/2 to VLAN 2 and GigabitEthernet 1/0/1 to VLAN 3. Configure the IP address 192.168.10.100/16 for the interface of VLAN 10.

```
<Switch> system-view
[Switch] vlan 2
[Switch-vlan2] port GigabitEthernet 1/0/2
[Switch-vlan2] quit
[Switch] vlan 3
[Switch-vlan3] port GigabitEthernet 1/0/1
[Switch-vlan3] quit
[Switch] vlan 10
[Switch-vlan10] supervlan
[Switch-vlan10] subvlan 2 3
[Switch-vlan10] quit
[Switch] interface vlan-interface 10
[Switch-Vlan-interface10] ip address 192.168.10.100 255.255.0.0
```

From Host A, ping Host B. The ping operation is unsuccessful because they are isolated at Layer 2.

Configure local proxy ARP to implement Layer 3 communication between sub-VLANs.

```
[Switch-Vlan-interface10] local-proxy-arp enable
```

From Host A, ping Host B. The ping operation is successful after the configuration.

Local proxy ARP configuration example in isolate-user-VLAN

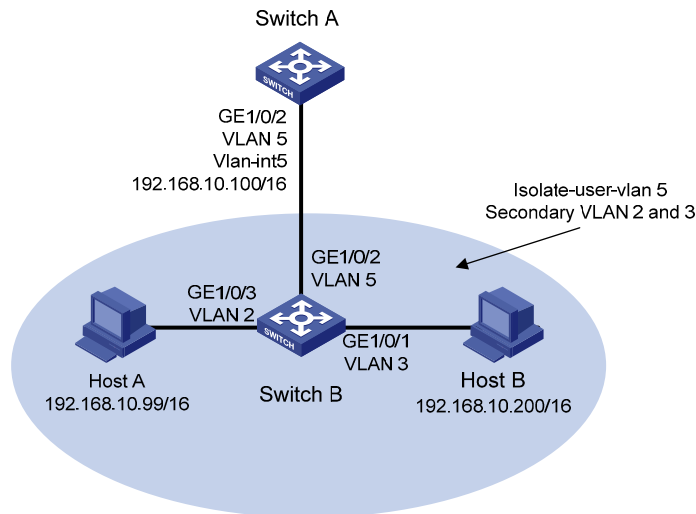
Network requirements

As shown in [Figure 11](#), Switch B is attached to Switch A. VLAN 5 on Switch B is an isolate-user-VLAN, which includes uplink port GigabitEthernet 1/0/2 and two secondary VLANs, VLAN 2 and VLAN 3. GigabitEthernet 1/0/3 belongs to VLAN 2, and GigabitEthernet 1/0/1 belongs to VLAN 3.

Host A belongs to VLAN 2 and connects to GigabitEthernet 1/0/3 of Switch B. Host B belongs to VLAN 3 and connects to GigabitEthernet 1/0/1 of Switch B.

As Host A and Host B belong to different secondary VLANs, they are isolated at Layer 2. Configure local proxy ARP on Switch A to implement Layer 3 communication between Host A and Host B.

Figure 11 Network diagram



Configuration procedure

1. Configure Switch B:

Create VLAN 2, VLAN 3, and VLAN 5 on Switch B. Add GigabitEthernet 1/0/3 to VLAN 2, GigabitEthernet 1/0/1 to VLAN 3, and GigabitEthernet 1/0/2 to VLAN 5. Configure VLAN 5 as the isolate-user-VLAN, and VLAN 2 and VLAN 3 as secondary VLANs. Configure the mappings between isolate-user-VLAN and the secondary VLANs.

```
<SwitchB> system-view
[SwitchB] vlan 2
[SwitchB-vlan2] port GigabitEthernet 1/0/3
[SwitchB-vlan2] quit
[SwitchB] vlan 3
[SwitchB-vlan3] port GigabitEthernet 1/0/1
[SwitchB-vlan3] quit
[SwitchB] vlan 5
[SwitchB-vlan5] port GigabitEthernet 1/0/2
[SwitchB-vlan5] isolate-user-vlan enable
[SwitchB-vlan5] quit
[SwitchB] interface GigabitEthernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] port isolate-user-vlan 5 promiscuous
[SwitchB-GigabitEthernet1/0/2] quit
[SwitchB] interface GigabitEthernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] port isolate-user-vlan host
[SwitchB-GigabitEthernet1/0/1] quit
[SwitchB] interface GigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] port isolate-user-vlan host
[SwitchB-GigabitEthernet1/0/3] quit
[SwitchB] isolate-user-vlan 5 secondary 2 3
```

2. Configure Switch A:

Create VLAN 5 and add GigabitEthernet 1/0/2 to it.

```
<SwitchA> system-view
[SwitchA] vlan 5
```

```
[SwitchA-vlan5] port GigabitEthernet 1/0/2
[SwitchA-vlan5] quit
[SwitchA] interface vlan-interface 5
[SwitchA-Vlan-interface5] ip address 192.168.10.100 255.255.0.0
```

From Host A, ping Host B. The ping operation is unsuccessful because they are isolated at Layer 2.

Configure local proxy ARP to implement Layer 3 communication between Host A and Host B.

```
[SwitchA-Vlan-interface5] local-proxy-arp enable
```

From Host A, ping Host B. The ping operation is successful after the configuration.

Configuring ARP snooping

Overview

The ARP snooping feature is used in Layer 2 switching networks. It creates ARP snooping entries using ARP packets.

If ARP snooping is enabled on a VLAN of a device, ARP packets received by the interfaces of the VLAN are redirected to the CPU. The CPU uses ARP packets to create ARP snooping entries comprising source IP and MAC addresses, VLAN and receiving port information.

The aging time and valid period of an ARP snooping entry are 25 minutes and 15 minutes, respectively. If an ARP snooping entry is not updated within 15 minutes, it becomes invalid and cannot be used. After that, if an ARP packet whose source IP and MAC addresses correspond with the entry is received, the entry becomes valid, and its age timer restarts. If the age timer of an ARP entry expires, the entry is removed.

If the ARP snooping device receives an ARP packet that has the same sender IP address as but a different sender MAC address from a valid ARP snooping entry, it considers that an attack occurs. An ARP snooping entry conflict occurs in this case. As a result, the ARP snooping entry becomes invalid and is removed after 25 minutes.

Configuration procedure

To enable ARP snooping for a VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
3. Enable ARP snooping.	arp-snooping enable	Disabled by default

Displaying and maintaining ARP snooping

Task	Command	Remarks
Display ARP snooping entries.	display arp-snooping [ip <i>ip-address</i> vlan <i>vlan-id</i>] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Remove ARP snooping entries.	reset arp-snooping [ip <i>ip-address</i> vlan <i>vlan-id</i>]	Available in user view

Configuring IP addressing

Only the HP 5500 EI switches support Layer 3 Ethernet port configuration.

This chapter describes IP addressing basic and manual IP address assignment for interfaces. Dynamic IP address assignment (BOOTP and DHCP) are beyond the scope of this chapter.

The term "interface" in this chapter refers to Layer 3 interfaces, including VLAN interfaces and route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).

Overview

This section describes the IP addressing basics.

IP addressing uses a 32-bit address to identify each host on a network. To make addresses easier to read, they are written in dotted decimal notation, each address being four octets in length. For example, address 000010000000000010000000100000001 in binary is written as 10.1.1.1.

IP address classes

Each IP address breaks down into two parts:

- **Net ID**—Identifies a network. The first several bits of a net ID, known as the class field or class bits, identify the class of the IP address.
- **Host ID**—Identifies a host on a network.

IP addresses are divided into five classes, shown in [Figure 12](#). The shaded areas represent the address class. The first three classes are widely used.

Figure 12 IP address classes

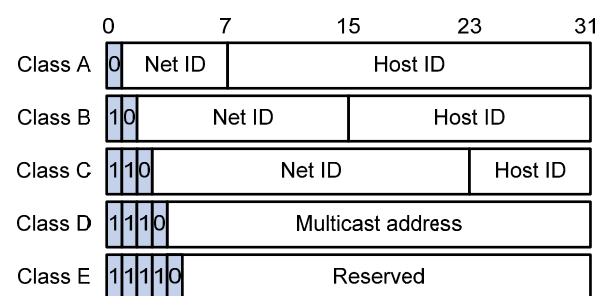


Table 1 IP address classes and ranges

Class	Address range	Remarks
A	0.0.0.0 to 127.255.255.255	The IP address 0.0.0.0 is used by a host at startup for temporary communication. This address is never a valid destination address. Addresses starting with 127 are reserved for loopback test. Packets destined to these addresses are processed locally as input packets rather than sent to the link.

Class	Address range	Remarks
B	128.0.0.0 to 191.255.255.255	N/A
C	192.0.0.0 to 223.255.255.255	N/A
D	224.0.0.0 to 239.255.255.255	Multicast addresses.
E	240.0.0.0 to 255.255.255.255	Reserved for future use except for the broadcast address 255.255.255.255.

Special IP addresses

The following IP addresses are for special use and cannot be used as host IP addresses.

- **IP address with an all-zero net ID**—Identifies a host on the local network. For example, IP address 0.0.0.16 indicates the host with a host ID of 16 on the local network.
- **IP address with an all-zero host ID**—Identifies a network.
- **IP address with an all-one host ID**—Identifies a directed broadcast address. For example, a packet with the destination address of 192.168.1.255 will be broadcast to all the hosts on the network 192.168.1.0.

Subnetting and masking

Subnetting divides a network down into smaller networks called subnets by using some bits of the host ID to create a subnet ID.

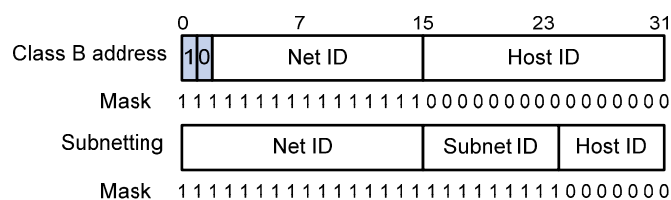
Masking identifies the boundary between the host ID and the combination of net ID and subnet ID. (When subnetting is not adopted, a mask identifies the boundary between the net ID and the host ID.)

Each subnet mask is made up of 32 bits that correspond to the bits in an IP address. In a subnet mask, consecutive ones represent the net ID and subnet ID, and consecutive zeros represent the host ID.

Before being subnetted, Class A, B, and C networks use the following default masks (also called natural masks): 255.0.0.0, 255.255.0.0, and 255.255.255.0 respectively.

Figure 13 shows how a Class B network is subnetted.

Figure 13 Subnetting a Class B network



Subnetting increases the number of addresses that cannot be assigned to hosts. After being subnetted, a network can accommodate fewer hosts.

For example, a Class B network without subnetting can accommodate 1022 more hosts than the same network subnetted into 512 subnets.

- **Without subnetting**—65,534 hosts ($2^{16} - 2$). (The two deducted addresses are the broadcast address, which has an all-one host ID, and the network address, which has an all-zero host ID.)
- **With subnetting**—Using the first 9 bits of the host-id for subnetting provides 512 (2^9) subnets. However, only 7 bits remain available for the host ID. This allows 126 ($2^7 - 2$) hosts in each subnet, a total of 64,512 hosts (512×126).

Assigning an IP address to an interface

You can assign an interface one primary address and multiple secondary addresses.

Generally, you only need to assign the primary address to an interface. In some cases, you need to assign secondary IP addresses to the interface. For example, if the interface connects to two subnets, to enable the device to communicate with all hosts on the LAN, you need to assign a primary IP address and a secondary IP address to the interface.

Configuration guidelines

Follow these guidelines when you assign an IP address to an interface:

- Each interface has only one primary IP address. A newly configured primary IP address overwrites the previous one.
- You cannot assign secondary IP addresses to an interface that obtains an IP address through BOOTP or DHCP.
- The primary and secondary IP addresses you assign to the interface can be located on the same network segment, but different interfaces on your device must reside on different network segments.
- You can manually assign an IP address to an interface, or configure the interface to obtain an IP address through BOOTP or DHCP. If you change the way an interface obtains an IP address, the new IP address overwrites the previous one.

Configuration procedure

To assign an IP address to an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Assign an IP address to the interface.	ip address <i>ip-address</i> { <i>mask-length</i> <i>mask</i> } [sub]	By default, no IP address is assigned to any interface.

Configuration example

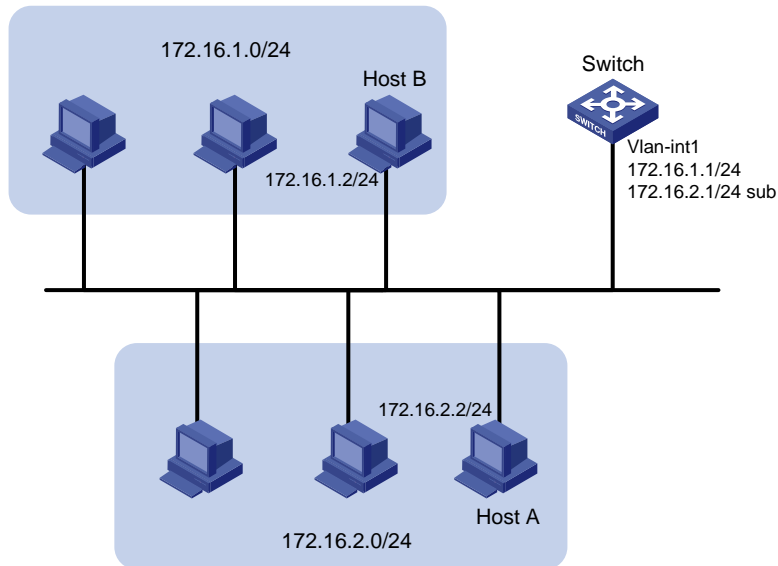
Network requirements

As shown in [Figure 14](#), a port in VLAN 1 on a switch is connected to a LAN comprising two segments: 172.16.1.0/24 and 172.16.2.0/24.

To enable the hosts on the two subnets to communicate with the external network through the switch, and to enable the hosts on the two subnets to communicate with each other:

- Assign a primary IP address and a secondary IP address to VLAN-interface 1 on the switch.
- Set the primary IP address of VLAN-interface 1 as the gateway address of the hosts on subnet 172.16.1.0/24, and the secondary IP address of VLAN-interface 1 as the gateway address of the hosts on subnet 172.16.2.0/24.

Figure 14 Network diagram



Configuration procedure

Assign a primary IP address and a secondary IP address to VLAN-interface 1.

```
<Switch> system-view
[Switch] interface vlan-interface 1
[Switch-Vlan-interface1] ip address 172.16.1.1 255.255.255.0
[Switch-Vlan-interface1] ip address 172.16.2.1 255.255.255.0 sub
```

Set the gateway address to 172.16.1.1 on the hosts attached to subnet 172.16.1.0/24, and to 172.16.2.1 on the hosts attached to subnet 172.16.2.0/24.

From the switch, ping a host on subnet 172.16.1.0/24 to verify the connectivity.

```
<Switch> ping 172.16.1.2
PING 172.16.1.2: 56 data bytes, press CTRL_C to break
  Reply from 172.16.1.2: bytes=56 Sequence=1 ttl=255 time=25 ms
  Reply from 172.16.1.2: bytes=56 Sequence=2 ttl=255 time=27 ms
  Reply from 172.16.1.2: bytes=56 Sequence=3 ttl=255 time=26 ms
  Reply from 172.16.1.2: bytes=56 Sequence=4 ttl=255 time=26 ms
  Reply from 172.16.1.2: bytes=56 Sequence=5 ttl=255 time=26 ms
```

```
--- 172.16.1.2 ping statistics ---
  5 packet(s) transmitted
  5 packet(s) received
  0.00% packet loss
  round-trip min/avg/max = 25/26/27 ms
```

The output shows that the switch can communicate with the hosts on subnet 172.16.1.0/24.

From the switch, ping a host on subnet 172.16.2.0/24 to verify the connectivity.

```
<Switch> ping 172.16.2.2
PING 172.16.2.2: 56 data bytes, press CTRL_C to break
  Reply from 172.16.2.2: bytes=56 Sequence=1 ttl=255 time=25 ms
  Reply from 172.16.2.2: bytes=56 Sequence=2 ttl=255 time=26 ms
  Reply from 172.16.2.2: bytes=56 Sequence=3 ttl=255 time=26 ms
  Reply from 172.16.2.2: bytes=56 Sequence=4 ttl=255 time=26 ms
  Reply from 172.16.2.2: bytes=56 Sequence=5 ttl=255 time=26 ms

--- 172.16.2.2 ping statistics ---
  5 packet(s) transmitted
  5 packet(s) received
  0.00% packet loss
  round-trip min/avg/max = 25/25/26 ms
```

The output shows that the switch can communicate with the hosts on subnet 172.16.2.0/24.

From a host on subnet 172.16.2.0/24, ping a host on subnet 172.16.1.0/24 to verify the connectivity. Host B can be successfully pinged from Host A.

Configuring IP unnumbered(only available on the HP 5500 EI)

Overview

Logically, to enable IP on an interface, you must assign this interface a unique IP address. Yet, you can borrow an IP address already configured on one of other interfaces on your device instead. This is called "IP unnumbered" and the interface borrowing the IP address is called "IP unnumbered interface".

You can use IP unnumbered to save IP addresses either when available IP addresses are inadequate or when an interface is brought up only for occasional use.

Configuration guidelines

Follow these guidelines when you configure IP unnumbered on an interface:

- An interface cannot borrow an IP address from an unnumbered interface.
- Multiple interfaces can use the same unnumbered IP address.
- If an interface has multiple IP addresses, only the primary IP address can be borrowed.
- The IP address of the borrowing interface varies with that of the borrowed interface. If an IP address is configured for the borrowed interface, the IP address of the borrowing interface is the same as that of the borrowed interface; if no IP address is configured for the borrowed interface, no IP address is assigned for the borrowing interface.

Configuration prerequisites

Assign a primary IP address to the interface from which you want to borrow the IP address. Alternatively, you may configure the interface to obtain one through BOOTP or DHCP.

Configuration procedure

To configure IP unnumbered on an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter tunnel interface view.	interface tunnel <i>number</i>	N/A
3. Specify the current interface to borrow the IP address of the specified interface.	ip address unnumbered interface <i>interface-type interface-number</i>	The interface does not borrow IP addresses from other interfaces by default.

Displaying and maintaining IP addressing

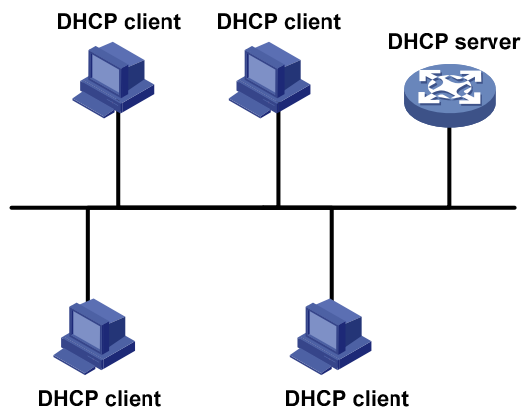
Task	Command	Remarks
Display IP configuration information for a specified Layer 3 interface or all Layer 3 interfaces.	display ip interface [<i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display brief IP configuration information for a specified Layer 3 interface or all Layer 3 interfaces.	display ip interface [<i>interface-type [interface-number]</i>] brief [{ begin exclude include } <i>regular-expression</i>]	Available in any view

DHCP overview

The Dynamic Host Configuration Protocol (DHCP) provides a framework to assign configuration information to network devices.

DHCP uses the client/server model.

Figure 15 A typical DHCP application



A DHCP client can obtain an IP address and other configuration parameters from a DHCP server on another subnet via a DHCP relay agent. For more information about the DHCP relay agent, see "[Configuring DHCP relay agent.](#)"

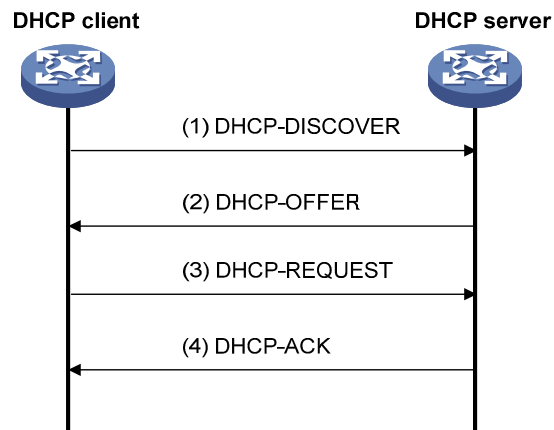
DHCP address allocation

DHCP supports the following mechanisms for IP address allocation.

- **Static allocation**—The network administrator assigns an IP address to a client like a WWW server, and DHCP conveys the assigned address to the client.
- **Automatic allocation**—DHCP assigns a permanent IP address to a client.
- **Dynamic allocation**—DHCP assigns an IP address to a client for a limited period of time, which is called a lease. Most DHCP clients obtain their addresses in this way.

Dynamic IP address allocation process

Figure 16 Dynamic IP address allocation process



1. The client broadcasts a DHCP-DISCOVER message to locate a DHCP server.
2. A DHCP server offers configuration parameters such as an IP address to the client, in a DHCP-OFFER message. The sending mode of the DHCP-OFFER is determined by the flag field in the DHCP-DISCOVER message. For related information, see "[DHCP message format](#)."
3. If several DHCP servers send offers to the client, the client accepts the first received offer, and broadcasts it in a DHCP-REQUEST message to formally request the IP address.
4. All DHCP servers receive the DHCP-REQUEST message, but only the server from which the client accepts the offered IP address returns either a DHCP-ACK message to the client, confirming that the IP address has been allocated to the client, or a DHCP-NAK message, denying the IP address allocation.

After the client receives the DHCP-ACK message, it broadcasts a gratuitous ARP packet to verify whether the IP address assigned by the server is already in use. If the client receives no response within the specified time, the client uses the assigned IP address. Otherwise, the client sends a DHCP-DECLINE message to the server to request an IP address again.

IP addresses offered by other DHCP servers are still assignable to other clients.

IP address lease extension

The IP address dynamically allocated by a DHCP server to a client has a lease. When the lease expires, the IP address is reclaimed by the DHCP server. To continue using the IP address, the client must extend the lease duration.

After half the lease duration, the DHCP client sends a DHCP-REQUEST unicast to the DHCP server to extend the lease. Depending on availability of the IP address, the DHCP server returns either a DHCP-ACK unicast confirming that the client's lease has been extended, or a DHCP-NAK unicast denying the request.

If the client receives no reply, it broadcasts another DHCP-REQUEST message for lease extension after 7/8 lease duration. Again, depending on availability of the IP address, the DHCP server returns either a DHCP-ACK unicast confirming that the client's lease has been extended, or a DHCP-NAK unicast denying the request.

DHCP message format

Figure 17 shows the DHCP message format, which is based on the BOOTP message format although DHCP uses some of the fields in significantly different ways. The numbers in parentheses indicate the size of each field in bytes.

Figure 17 DHCP message format

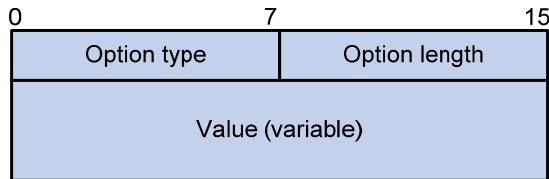


- **op**—Message type defined in option field. 1 = REQUEST, 2 = REPLY
- **htype, hlen**—Hardware address type and length of a DHCP client.
- **hops**—Number of relay agents a request message traveled.
- **xid**—Transaction ID, a random number chosen by the client to identify an IP address allocation.
- **secs**—Filled in by the client, the number of seconds elapsed since the client began address acquisition or renewal process. Currently this field is reserved and set to 0.
- **flags**—The leftmost bit is defined as the BROADCAST (B) flag. If this flag is set to 0, the DHCP server sent a reply back by unicast; if this flag is set to 1, the DHCP server sent a reply back by broadcast. The remaining bits of the flags field are reserved for future use.
- **ciaddr**—Client IP address if the client has an IP address that is valid and usable; otherwise, set to zero.
- **yiaddr**—'Your' (client) IP address, assigned by the server.
- **siaddr**—Server IP address, from which the client obtained configuration parameters.
- **giaddr**—(Gateway) IP address of the first relay agent a request message traveled.
- **chaddr**—Client hardware address.
- **sname**—Server host name, from which the client obtained configuration parameters.
- **file**—Bootfile name and path information, defined by the server to the client.
- **options**—Optional parameters field that is variable in length, which includes the message type, lease duration, subnet mask, domain name server IP address, WINS IP address, and other information.

DHCP options

DHCP uses the same message format as BOOTP, but DHCP uses the Option field to carry information for dynamic address allocation and to provide additional configuration information to clients.

Figure 18 DHCP option format



Common DHCP options

The following are common DHCP options:

- **Option 3**—Router option. It specifies the gateway address.
- **Option 6**—DNS server option. It specifies the DNS server’s IP address.
- **Option 33**—Static route option. It specifies a list of classful static routes (the destination addresses in these static routes are classful) that a client should add into its routing table. If both Option 33 and Option 121 exist, Option 33 is ignored.
- **Option 51**—IP address lease option.
- **Option 53**—DHCP message type option. It identifies the type of the DHCP message.
- **Option 55**—Parameter request list option. It is used by a DHCP client to request specified configuration parameters. The option contains values that correspond to the parameters requested by the client.
- **Option 60**—Vendor class identifier option. It is used by a DHCP client to identify its vendor, and by a DHCP server to distinguish DHCP clients by vendor class and assign specific IP addresses for the DHCP clients.
- **Option 66**—TFTP server name option. It specifies a TFTP server to be assigned to the client.
- **Option 67**—Bootfile name option. It specifies the bootfile name to be assigned to the client.
- **Option 121**—Classless route option. It specifies a list of classless static routes (the destination addresses in these static routes are classless) that the requesting client should add to its routing table. If both Option 33 and Option 121 exist, Option 33 is ignored.
- **Option 150**—TFTP server IP address option. It specifies the TFTP server IP address to be assigned to the client.

For more information about DHCP options, see RFC 2132 and RFC 3442.

Custom options

Some options, such as Option 43, Option 82, and Option 184, have no unified definitions in RFC 2132.

Vendor-specific option (Option 43)

DHCP servers and clients use Option 43 to exchange vendor-specific configuration information.

The DHCP client can obtain the following information through Option 43:

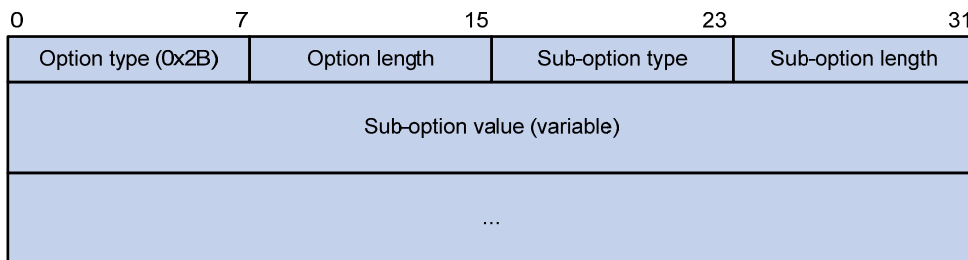
- Auto-Configuration Server (ACS) parameters, including the ACS URL, username, and password.

- Service provider identifier, which is acquired by the Customer Premises Equipment (CPE) from the DHCP server and sent to the ACS for selecting vender-specific configurations and parameters.
- Preboot Execution Environment (PXE) server address, which is used to obtain the bootfile or other control information from the PXE server.

1. Format of Option 43

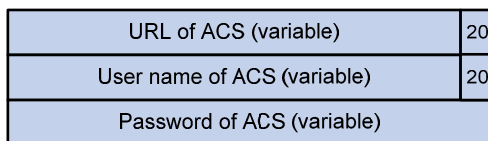
Network configuration parameters are carried in different sub-options of Option 43 as shown in Figure 19.

Figure 19 Option 43 format



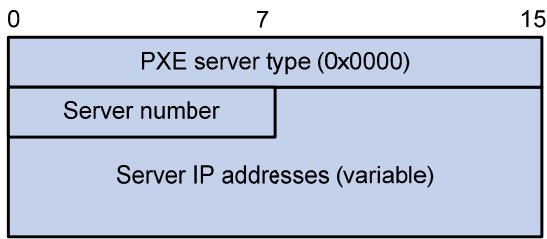
- **Sub-option type**—Type of a sub-option. The field value can be 0x01 (an ACS parameter sub-option), 0x02 (a service provider identifier sub-option), or 0x80 (a PXE server address sub-option).
 - **Sub-option length**—Length of a sub-option excluding the sub-option type and sub-option length fields.
 - **Sub-option value**—Value of a sub-option. The value format varies with sub-options.
2. Format of the sub-option value field of Option 43
- As shown in Figure 20, the value field of the ACS parameter sub-option contains variable ACS URL, ACS username, and ACS password separated by spaces (0x20):

Figure 20 ACS parameter sub-option value field



- The value field of the service provider identifier sub-option contains the service provider identifier.
- Figure 21 shows the format of the value field of the PXE server address sub-option. The value of the PXE server type can only be 0. The server number field indicates the number of PXE servers contained in the sub-option. The server IP addresses field contains the IP addresses of the PXE servers.

Figure 21 PXE server address sub-option value field



Relay agent option (Option 82)

Option 82 is the relay agent option in the option field of the DHCP message. It records the location information of the DHCP client. When a DHCP relay agent or DHCP snooping device receives a client's request, it adds Option 82 to the request message and sends it to the server.

The administrator can locate the DHCP client to further implement security control and accounting. The Option 82 supporting server can also use such information to define individual assignment policies of IP address and other parameters for the clients.

Option 82 involves at most 255 sub-options. At least one sub-option must be defined. The DHCP relay agent supports three sub-options: sub-option 1 (Circuit ID), sub-option 2 (Remote ID) and sub-option 9 (private padding format).

Option 82 has no unified definition. Its padding formats vary with vendors.

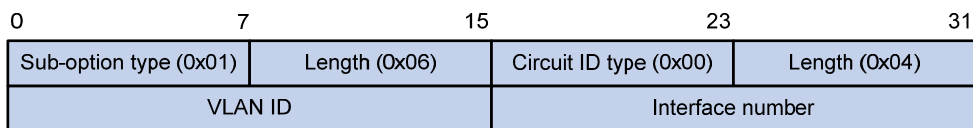
There are two methods for configuring Option 82:

- **User-defined method**—Manually specify the content of Option 82.
- **Non-user-defined method**—Pad Option 82 in the default normal format, verbose format, private format, or standard format.

If you choose normal format and verbose format, you can specify the code type for the sub-options as ASCII or HEX.

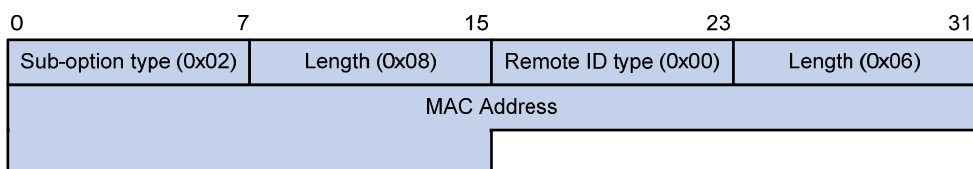
- Normal padding format
 - **Sub-option 1**—Contains the VLAN ID and interface number of the interface that received the client's request. The value of the sub-option type is 1, and that of the circuit ID type is 0.

Figure 22 Sub-option 1 in normal padding format



- **Sub-option 2**—Contains the MAC address of the DHCP relay agent interface or the MAC address of the DHCP snooping device that received the client's request. The value of the sub-option type is 2, and that of the remote ID type is 0.

Figure 23 Sub-option 2 in normal padding format



- Verbose padding format
 - **Sub-option 1**—Contains the user-specified access node identifier (ID of the device that adds Option 82 in DHCP messages), and the type, number, and VLAN ID of the interface that received the client’s request. The VLAN ID field has a fixed length of 2 bytes. All the other padding contents of sub-option 1 are length variable. See [Figure 24](#).

Figure 24 Sub-option 1 in verbose padding format

Sub-option type (0x01)	Length	Node identifier
Interface type		Interface number
VLAN ID		

- **Sub-option 2**—Contains the MAC address of the DHCP relay agent interface or the MAC address of the DHCP snooping device that received the client’s request. It has the same format as that in normal padding format. See [Figure 23](#).
- Private padding format
 - **Sub-option 1**—Contains the VLAN ID of the interface that received the client’s request, module (subcard number of the receiving port) and port (port number of the receiving port). The value of the sub-option type is 1.

Figure 25 Sub-option 1 in private padding format

0	7	15	23	31
Sub-option type (0x01)	Length (0x04)	VLAN ID		
Module		Port		

- **Sub-option 2**—Contains the MAC address of the DHCP relay agent interface or the MAC address of the DHCP snooping device that received the client’s request. The value of the sub-option type is 2.

Figure 26 Sub-option 2 in private padding format

0	7	15	23	31
Sub-option type (0x02)	Length (0x06)	MAC Address		

- **Sub-option 9**—Contains the sysname and the primary IP address of the Loopback0 interface. The value of the sub-option type is 9.

Figure 27 Sub-option 9 in private padding format

0	7	15	23	31
Sub-option type (0x09)	Length	Enterprise Number		
Enterprise Number		Information Length	Index (0x01)	
Index (0x00)	Index (0x02)	Length	Sysname	
Sysname		Index (0x03)	Index (0x04)	
LoopBack0 IP				

- Standard padding format

- **Sub-option 1**—Contains the VLAN ID of the interface that received the client’s request, module (subcard number of the receiving port) and port (port number of the receiving port). The value of the sub-option type is 1, and the value of the circuit ID type is 0.

Figure 28 Sub-option 1 in standard padding format

0	7	15	23	31
Sub-option type (0x01)		Length (0x06)		Circuit ID type (0x00)
VLAN ID		Module		Port

- **Sub-option 2**—Contains the MAC address of the DHCP relay agent interface or the MAC address of the DHCP snooping device that received the client’s request. It has the same format as sub-option 2 in normal padding format. See [Figure 23](#).

Option 184

Option 184 is a reserved option, and parameters in the option can be defined as needed. The device supports Option 184 carrying voice related parameters, so a DHCP client with voice functions can get an IP address along with specified voice parameters from the DHCP server.

Option 184 involves the following sub-options:

- **Sub-option 1**—IP address of the primary network calling processor, which serves as the network calling control source and provides program downloads.
- **Sub-option 2**—IP address of the backup network calling processor. DHCP clients contact the backup when the primary is unreachable.
- **Sub-option 3**—Voice VLAN ID and the result whether or not DHCP clients take this ID as the voice VLAN.
- **Sub-option 4**—Failover route that specifies the destination IP address and the called number. A Session Initiation Protocol (SIP) user uses this IP address and number to reach another SIP user when both the primary and backup calling processors are unreachable.

You must define sub-option 1 to make other sub-options take effect.

Protocols and standards

- RFC 2131, *Dynamic Host Configuration Protocol*
- RFC 2132, *DHCP Options and BOOTP Vendor Extensions*
- RFC 1542, *Clarifications and Extensions for the Bootstrap Protocol*
- RFC 3046, *DHCP Relay Agent Information Option*
- RFC 3442, *The Classless Static Route Option for Dynamic Host Configuration Protocol (DHCP) version 4*

Configuring DHCP server

The term "interface" in the DHCP features refers to Layer 3 interfaces, including VLAN interfaces and route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see the *Layer 2—LAN Switching Configuration Guide*).

Only HP 5500 EI switches support Layer 3 Ethernet port configuration.

Overview

The DHCP server is well suited to networks where:

- Manual configuration and centralized management are difficult to implement.
- Many hosts need to acquire IP addresses dynamically. This may be because the number of hosts exceeds the number of assignable IP addresses, so it is impossible to assign a fixed IP address to each host. For example, an ISP has a limited number of host addresses.
- A few hosts need fixed IP addresses.

In addition to assigning IP addresses to DHCP clients on public networks, a multi-VPN-instance customer edge (MCE) serving as the DHCP server can also assign IP addresses to DHCP clients on private networks. Note that the IP address ranges of public and private networks or those of private networks on the DHCP server cannot overlap each other. For more information about MCE, see *Layer 3—IP Routing Configuration Guide*.

DHCP address pool

Address pool types

DHCP address pools include common and extended address pools.

- **Common address pool**—Supports both static binding and dynamic allocation.
- **Extended address pool**—Supports only dynamic allocation.

Common address pool structure

The common address pool database is organized as a tree. The root of the tree is the address pool for natural networks, branches are address pools for subnets, and leaves are addresses statically bound to clients. For the same level address pools, a previously configured pool has a higher selection priority than a new one.

At the very beginning, subnets inherit network parameters and clients inherit subnet parameters. Therefore, common parameters, for example a DNS server address, should be configured at the highest (network or subnet) level of the tree. IP address lease durations are not inherited.

The new configuration at the higher level (parent) of the tree will be:

- Inherited if the lower level (child) has no such configuration.
- Overridden if the lower level (child) has such configuration.

NOTE:

The extended address pools on a DHCP server are independent of each other and no inheritance relationship exists among them.

Principles for selecting an address pool

The DHCP server observes the following principles to select an address pool when assigning an IP address to a client:

1. If there is an address pool where an IP address is statically bound to the MAC address or ID of the client, the DHCP server will select this address pool and assign the statically bound IP address to the client. For the configuration of this address pool, see "[Configuring static address allocation](#)."
2. If the receiving interface has an extended address pool referenced, the DHCP server will assign an IP address from this address pool. If no IP address is available in the address pool, the DHCP server will fail to assign an address to the client. For the configuration of such an address pool, see "[Configuring dynamic address allocation for an extended address pool](#)."
3. Otherwise, the DHCP server will select the smallest common address pool that contains the IP address of the receiving interface (if the client and the server reside on the same subnet), or the smallest common address pool that contains the IP address specified in the giaddr field of the client's request (if a DHCP relay agent is in-between). If no IP address is available in the address pool, the DHCP server will fail to assign an address to the client because it cannot assign an IP address from the parent address pool to the client. For the configuration of such an address pool, see "[Configuring dynamic address allocation](#)."

For example, two common address pools, 1.1.1.0/24 and 1.1.1.0/25, are configured on the DHCP server. If the IP address of the interface receiving DHCP requests is 1.1.1.1/25, the DHCP server will select IP addresses for clients from address pool 1.1.1.0/25. If no IP address is available in the address pool, the DHCP server will fail to assign addresses to clients. If the IP address of the interface receiving DHCP requests is 1.1.1.130/25, the DHCP server will select IP addresses for clients from the 1.1.1.0/24 address pool.

NOTE:

To avoid wrong IP address allocation, keep the IP addresses for dynamic allocation within the subnet where the interface of the DHCP server or DHCP relay agent resides.

IP address allocation sequence

A DHCP server assigns an IP address to a client according to the following sequence:

1. The IP address statically bound to the client's MAC address or ID.
2. The IP address that was ever assigned to the client.
3. The IP address designated by the Option 50 field in a DHCP-DISCOVER message. Option 50 is the requested IP address field in DHCP-DISCOVER messages. It is padded by the client to specify the IP address that the client wants to obtain. The contents to be padded depend on the client.
4. The first assignable IP address found in an extended or common address pool.
5. The IP address that was a conflict or passed its lease duration.

If no IP address is assignable, the server will not respond.

DHCP server configuration task list

Task	Remarks
Configuring an address pool for the DHCP server	Required.
Enabling DHCP	Required.
Enabling the DHCP server on an interface	Required.
Applying an extended address pool on an interface	Required by the extended address pool configuration. When configuring a common address pool, ignore this task.
Configuring the DHCP server security functions	Optional.
Enabling client offline detection	Optional.
Enabling handling of Option 82	Optional.
Specifying a server's IP address for the DHCP client	Optional.
Specifying the threshold for sending trap messages	Optional.
Setting the DSCP value for DHCP packets	Optional.

Configuring an address pool for the DHCP server Configuration task list

Task	Remarks			
Creating a DHCP address pool	Required.			
Configuring address allocation mode for a common address pool	<table border="0"> <tr> <td>Configuring static address allocation</td> <td rowspan="2">Required to configure either of the two for the common address pool configuration.</td> </tr> <tr> <td>Configuring dynamic address allocation</td> </tr> </table>	Configuring static address allocation	Required to configure either of the two for the common address pool configuration.	Configuring dynamic address allocation
Configuring static address allocation	Required to configure either of the two for the common address pool configuration.			
Configuring dynamic address allocation				
Configuring dynamic address allocation for an extended address pool	Required for the extended address pool configuration.			
Configuring a domain name suffix for the client				
Configuring DNS servers for the client				
Configuring WINS servers and NetBIOS node type for the client				
Configuring BIMS server information for the client				
Configuring gateways for the client	Optional.			
Configuring Option 184 parameters for the client with voice service				
Configuring the TFTP server and bootfile name for the client				
Specifying a server's IP address for the DHCP client				
Configuring self-defined DHCP options				

Creating a DHCP address pool

When creating a DHCP address pool, specify it as a common address pool or an extended address pool.

A common address pool and an extended address pool are different in address allocation mode configuration. Configurations of other parameters (such as the domain name suffix and DNS server address) for them are the same.

To create a DHCP address pool:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a DHCP address pool and enter its view.	dhcp server ip-pool <i>pool-name</i> [extended]	No DHCP address pool is created by default.

Configuring address allocation mode for a common address pool



IMPORTANT:

You can configure either a static binding or dynamic address allocation for a common address pool, but not both.

You need to specify a subnet for dynamic address allocation. A static binding is a special address pool containing only one IP address.

Configuring static address allocation

Some DHCP clients, such as a WWW server, need fixed IP addresses. To provide a fixed IP address, you can create a static binding of a client's MAC address or client ID to an IP address in the DHCP address pool. A static binding is a special address pool containing only one IP address.

When the client with that MAC address or client ID requests an IP address, the DHCP server will assign the IP address from the binding to the client.

Follow these guidelines when you configure a static binding in a common address pool:

- Use the `static-bind ip-address` command together with `static-bind mac-address` or `static-bind client-identifier` to accomplish a static binding configuration.
- In a DHCP address pool, if you execute the `static-bind mac-address` command before the `static-bind client-identifier` command, the latter will overwrite the former and vice versa.
- If you use the `static-bind ip-address`, `static-bind mac-address`, or `static-bind client-identifier` command repeatedly in the DHCP address pool, the new configuration will overwrite the previous one.
- The IP address of the static binding cannot be an interface address of the DHCP server. Otherwise, an IP address conflict may occur and the bound client cannot obtain an IP address correctly.
- The ID of the static binding must be identical to the ID displayed by using the `display dhcp client verbose` command on the client. Otherwise, the client cannot obtain an IP address.
- When the device serves as a DHCP client or BOOTP client, you must bind the DHCP client's ID to an IP address, or bind the BOOTP client's MAC address to an IP address on the DHCP server; otherwise, the DHCP or BOOTP client cannot obtain a static IP address.

- If the interfaces on a DHCP client share the same MAC address, you must specify the client ID, rather than MAC address, in a static binding to identify the requesting interface; otherwise, the client may fail to obtain an IP address.

To configure a static binding in a common address pool:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter common address pool view.	dhcp server ip-pool <i>pool-name</i>	N/A
3. Specify the IP address.	static-bind ip-address <i>ip-address</i> [<i>mask-length</i> mask <i>mask</i>]	No IP addresses are statically bound by default.
4. Specify the MAC address or client ID.	<ul style="list-style-type: none"> • Specify the MAC address: static-bind mac-address <i>mac-address</i> • Specify the client ID: static-bind client-identifier <i>client-identifier</i> 	Use at least one command. Neither is bound statically by default.
5. Specify the lease duration for the IP address.	expired { day <i>day</i> [hour <i>hour</i> [minute <i>minute</i> [second <i>second</i>]]] } unlimited }	Optional. By default, the lease duration of the IP address is unlimited.

Configuring dynamic address allocation

For dynamic address allocation, you must configure a DHCP address pool. For each address pool, you must specify one and only one address range, and the lease duration. A DHCP address pool can have only one lease duration.

To avoid address conflicts, configure the DHCP server to exclude IP addresses used by the gateway or FTP server from dynamic allocation.

Follow these guidelines when you configure dynamic address allocation for a common address pool:

- In common address pool view, using the **network** or **network ip range** command repeatedly overwrites the previous configuration.
- After you exclude IP addresses from automatic allocation by using the **dhcp server forbidden-ip** command, neither a common address pool nor an extended address pool can assign these IP addresses through dynamic address allocation.
- Using the **dhcp server forbidden-ip** command repeatedly can exclude multiple IP address ranges from allocation.

To configure dynamic address allocation for a common address pool:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter common address pool view.	dhcp server ip-pool <i>pool-name</i>	N/A
3. Specify a subnet.	network <i>network-address</i> [<i>mask-length</i> mask <i>mask</i>]	Not specified by default.

Step	Command	Remarks
4. Specify the IP address range on the subnet for dynamic allocation.	network ip range <i>min-address max-address</i>	Optional. Not specified by default.
5. Specify the address lease duration.	expired { day <i>day</i> [hour <i>hour</i> [minute <i>minute</i>] [second <i>second</i>]] unlimited }	Optional. One day by default.
6. Return to system view.	quit	N/A
7. Exclude IP addresses from automatic allocation.	dhcp server forbidden-ip <i>low-ip-address [high-ip-address]</i>	Optional. Except IP addresses of the DHCP server interfaces, all addresses in the DHCP address pool are assignable by default.

Configuring dynamic address allocation for an extended address pool

After the assignable IP address range and the mask are specified, the address pool becomes valid.

Extended address pools support dynamic address allocation only. Excluded IP addresses specified with the **forbidden-ip** command in DHCP address pool view are not assignable in the current extended address pool, but are assignable in other address pools.

To configure dynamic address allocation for an extended address pool:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter extended address pool view.	dhcp server ip-pool <i>pool-name</i> extended	N/A
3. Specify the IP address range.	network ip range <i>min-address max-address</i>	Not specified by default.
4. Specify the IP address mask.	network mask <i>mask</i>	Not specified by default.
5. Specify the IP address range for the DHCP clients of a specified vendor.	vendor-class-identifier <i>hex-string</i> <1-255> ip range <i>min-address max-address</i>	Optional. Not configured by default.
6. Specify the address lease duration.	expired { day <i>day</i> [hour <i>hour</i> [minute <i>minute</i> [second <i>second</i>]]] unlimited }	Optional. One day by default.
7. Exclude IP addresses from dynamic allocation.	forbidden-ip <i>ip-address</i> <1-8>	Optional. Except IP addresses of the DHCP server interfaces, all addresses in the DHCP address pool are assignable by default.

Configuring a domain name suffix for the client

You can specify a domain name suffix in each DHCP address pool on the DHCP server to provide the clients with the domain name suffix. With this suffix assigned, the client only needs to input part of a domain name, and the system will add the domain name suffix for name resolution. For more information about DNS, see "Configuring IPv4 DNS."

To configure a domain name suffix in the DHCP address pool:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter DHCP address pool view.	dhcp server ip-pool <i>pool-name</i> [extended]	N/A
3. Specify a domain name suffix.	domain-name <i>domain-name</i>	Not specified by default

Configuring DNS servers for the client

A DHCP client contacts a Domain Name System (DNS) server to resolve names. You can specify up to eight DNS servers in the DHCP address pool.

To configure DNS servers in the DHCP address pool:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter DHCP address pool view.	dhcp server ip-pool <i>pool-name</i> [extended]	N/A
3. Specify DNS servers.	dns-list <i>ip-address&<1-8></i>	Not specified by default

Configuring WINS servers and NetBIOS node type for the client

A Microsoft DHCP client using NetBIOS protocol contacts a Windows Internet Naming Service (WINS) server for name resolution. Therefore, the DHCP server should assign a WINS server address when assigning an IP address to the client.

You can specify up to eight WINS servers in a DHCP address pool.

You must also specify a NetBIOS node type in a DHCP address pool. There are four NetBIOS node types:

- **b (broadcast)-node**—A b-node client sends the destination name in a broadcast message. The destination returns its IP address to the client after receiving the message.
- **p (peer-to-peer)-node**—A p-node client sends the destination name in a unicast message to the WINS server, and the WINS server returns the destination IP address.
- **m (mixed)-node**—An m-node client broadcasts the destination name. If it receives no response, it unicasts the destination name to the WINS server to get the destination IP address.
- **h (hybrid)-node**—An h-node client unicasts the destination name to the WINS server. If it receives no response, it broadcasts the destination name to get the destination IP address.

To configure WINS servers and NetBIOS node type in the DHCP address pool:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter DHCP address pool view.	dhcp server ip-pool <i>pool-name</i> [extended]	N/A
3. Specify WINS server IP addresses.	nbns-list <i>ip-address</i> &<1-8>	Optional for b-node. No address is specified by default.
4. Specify the NetBIOS node type.	netbios-type { b-node h-node m-node p-node }	Not specified by default.

Configuring BIMS server information for the client

The DHCP server must provide DHCP clients with the branch intelligent management system (BIMS) server IP address, port number, shared key from the DHCP address pool, to enable DHCP clients to perform regular software update and backup by using configuration files obtained from a BIMS server.

To configure the BIMS server IP address, port number, and shared key in the DHCP address pool:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter DHCP address pool view.	dhcp server ip-pool <i>pool-name</i> [extended]	N/A
3. Specify the BIMS server IP address, port number, and shared key.	bims-server ip <i>ip-address</i> [port <i>port-number</i>] sharekey [cipher simple] <i>key</i>	Not specified by default

Configuring gateways for the client

You can specify up to eight gateways in a DHCP address pool.

To configure the gateways in the DHCP address pool:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter DHCP address pool view.	dhcp server ip-pool <i>pool-name</i> [extended]	N/A
3. Specify gateways.	gateway-list <i>ip-address</i> &<1-8>	No gateway is specified by default.

Configuring Option 184 parameters for the client with voice service

To assign voice calling parameters along with an IP address to DHCP clients with voice service, you must configure Option 184 on the DHCP server. For more information about Option 184, see "[DHCP overview](#)."

To configure option 184 parameters in the DHCP address pool:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter DHCP address pool view.	dhcp server ip-pool <i>pool-name</i> [extended]	N/A
3. Specify the IP address of the primary network calling processor.	voice-config ncp-ip <i>ip-address</i>	Not specified by default. After you configure this command, the other Option 184 parameters take effect.
4. Specify the IP address of the backup network calling processor.	voice-config as-ip <i>ip-address</i>	Optional. Not specified by default.
5. Configure the voice VLAN.	voice-config voice-vlan <i>vlan-id</i> { disable enable }	Optional. Not configured by default.
6. Specify the failover IP address and dialer string.	voice-config fail-over <i>ip-address dialer-string</i>	Optional. No failover IP address or dialer string is specified by default.

Configuring the TFTP server and bootfile name for the client

For the DHCP server to support client auto-configuration, you must specify the IP address or name of a TFTP server and the bootfile name in the DHCP address pool. You do not need to perform any configuration on the DHCP client.

The DHCP client uses these parameters to contact the TFTP server and request the configuration file used for system initialization.

1. When a switch starts up without loading any configuration file, the system sets an active interface (such as the interface of the default VLAN) as the DHCP client to request from the DHCP server for parameters, such as an IP address and name of a TFTP server, and the bootfile name.
2. After getting related parameters, the DHCP client will send a TFTP request to obtain the configuration file from the specified TFTP server for system initialization. If the client cannot get such parameters, it will perform system initialization without loading any configuration file.

To configure the IP address and name of the TFTP server and the bootfile name in the DHCP address pool:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter DHCP address pool view.	dhcp server ip-pool <i>pool-name</i> [extended]	N/A
3. Specify the IP address or name of the TFTP server.	<ul style="list-style-type: none"> Specify the TFTP server: tftp-server ip-address <i>ip-address</i> Specify the name of the TFTP server: tftp-server domain-name <i>domain-name</i> 	Use either command. Not specified by default.
4. Specify the bootfile name.	bootfile-name <i>bootfile-name</i>	Not specified by default.

Specifying a server's IP address for the DHCP client

Some DHCP clients need to obtain configuration information from a server, such as a TFTP server. You can specify the IP address of that server in each address pool of the DHCP server. The DHCP server sends the server's IP address to DHCP clients along with other configuration information.

To specify the IP address of a server:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter DHCP address pool view.	dhcp server ip-pool <i>pool-name</i> [extended]	N/A
3. Specify the IP address of a server.	next-server <i>ip-address</i>	Not specified by default

Configuring self-defined DHCP options

△ CAUTION:

Be cautious when configuring self-defined DHCP options because such configuration may affect the DHCP operation process.

By configuring self-defined DHCP options, you can

- Define new DHCP options. New configuration options will come out with DHCP development. To support these new options, you can add them into the attribute list of the DHCP server.
- Define existing DHCP options. Vendors use Option 43 to define options that have no unified definitions in RFC 2132. The self-defined DHCP option enables DHCP clients to obtain vendor-specific information.
- Extend existing DHCP options. When the current DHCP options cannot meet the customers' requirements (for example, you cannot use the **dns-list** command to configure more than eight DNS server addresses), you can configure a self-defined option for extension.

To configure a self-defined DHCP option in the DHCP address pool:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter DHCP address pool view.	dhcp server ip-pool <i>pool-name</i> [extended]	N/A

Step	Command	Remarks
3. Configure a self-defined DHCP option.	option code { ascii <i>ascii-string</i> hex <i>hex-string</i> &<1-16> ip-address <i>ip-address</i> &<1-8> }	No DHCP option is configured by default.

Table 2 Description of common options

Option	Option name	Corresponding command	Command parameter
3	Router Option	gateway-list	ip-address
6	Domain Name Server Option	dns-list	ip-address
15	Domain Name	domain-name	ascii
44	NetBIOS over TCP/IP Name Server Option	nbns-list	ip-address
46	NetBIOS over TCP/IP Node Type Option	netbios-type	hex
66	TFTP server name	tftp-server	ascii
67	Bootfile name	bootfile-name	ascii
43	Vendor Specific Information	N/A	hex

Enabling DHCP

Enable DHCP before performing other configurations.

To enable DHCP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable DHCP.	dhcp enable	Disabled by default

Enabling the DHCP server on an interface

With the DHCP server enabled on an interface, upon receiving a client's request, the DHCP server will assign an IP address from its address pool to the DHCP client.

Configuration guidelines

Follow these guidelines when you enable the DHCP server on an interface:

- If a DHCP relay agent exists between the DHCP server and client, the DHCP server, regardless of whether the **subaddress** keyword is used, will select an IP address from the address pool containing the primary IP address of the DHCP relay agent's interface (connected to the client) for a requesting client.
- When the DHCP server and client are on the same subnet:
 - With the keyword **subaddress** specified, the DHCP server will preferably assign an IP address from an address pool that resides on the same subnet as the primary IP address of the server

interface (connecting to the client). If the address pool contains no assignable IP address, the server assigns an IP address from an address pool that resides on the same subnet as the secondary IP addresses of the server interface. If the interface has multiple secondary IP addresses, each address pool is tried in turn for address allocation. If the interface has no secondary IP addresses, the server is unable to assign an IP address to the client.

- Without the keyword **subaddress** specified, the DHCP server can only assign an IP address from the address pool that resides on the same subnet as the primary IP address of the server interface.

Configuration procedure

To enable the DHCP server on an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable the DHCP server on an interface.	dhcp select server global-pool [subaddress]	Optional. Enabled by default.

Applying an extended address pool on an interface

After you create an extended address pool and apply it on an interface, the DHCP server, upon receiving a client's request on the interface, attempts to assign the client the statically bound IP address first and then an IP address from the specified address pool. If no IP address is available in this address pool, address allocation fails, and the DHCP server will not assign the client any IP address from other address pools.

Only an extended address pool can be applied on the interface. The address pool to be referenced must already exist.

To apply an extended address pool on an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Apply an extended address pool on the interface.	dhcp server apply ip-pool <i>pool-name</i>	Optional. By default, the DHCP server has no extended address pool applied on its interface, and assigns an IP address from a common address pool to a requesting client.

Configuring the DHCP server security functions

Configuration prerequisites

Before you configure the DHCP server security functions, complete the following tasks on the DHCP server:

- Enable DHCP.
- Configure the DHCP address pool.

Enabling unauthorized DHCP server detection

Unauthorized DHCP servers on a network may assign wrong IP addresses to DHCP clients.

With unauthorized DHCP server detection enabled, the DHCP server checks whether a DHCP request contains Option 54 (Server Identifier Option). If yes, the DHCP server records the IP address of each detected DHCP server that assigned an IP address to a requesting DHCP client in the option, and records the receiving interface. The administrator can use this information to check for unauthorized DHCP servers.

With the unauthorized DHCP server detection enabled, the switch logs each detected DHCP server once. The administrator can use the log information to find unauthorized DHCP servers.

To enable unauthorized DHCP server detection:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable unauthorized DHCP server detection.	dhcp server detect	Disabled by default

Configuring IP address conflict detection

With IP address conflict detection enabled, before assigning an IP address, the DHCP server pings that IP address by using ICMP. If the server receives a response within the specified period, it selects and pings another IP address. If it receives no response, the server continues to ping the IP address until the specified number of ping packets are sent. If still no response is received, the server assigns the IP address to the requesting client. (The DHCP client probes the IP address by sending gratuitous ARP packets.)

To configure IP address conflict detection:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Specify the number of ping packets.	dhcp server ping packets <i>number</i>	Optional. One ping packet by default. The value 0 indicates that no ping operation is performed.

Step	Command	Remarks
3. Configure a timeout waiting for ping responses.	dhcp server ping timeout <i>milliseconds</i>	Optional. 500 ms by default. The value 0 indicates that no ping operation is performed.

Enabling client offline detection

With this feature enabled, the DHCP server considers a DHCP client goes offline when the ARP entry for the client ages out. In addition, it removes the client's IP-to-MAC binding entry.

Removing an ARP entry manually does not remove the corresponding client's IP-to-MAC binding.

To enable offline detection:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable offline detection.	dhcp server client-detect enable	Disabled by default

Enabling handling of Option 82

With Option 82 handling enabled, when the DHCP server receives a request with Option 82, it adds Option 82 into the response.

If the server is configured to ignore Option 82, it will assign an IP address to the client without adding Option 82 in the response message.

Configuration prerequisites

Before you enable Option 82 handling, complete the following tasks:

- **Configure the DHCP server**—Enable DHCP and configure the DHCP address pool.
- **Configure the relay agent or the device enabled with DHCP snooping**—For more information, see "[Configuring DHCP relay agent](#)" and "[Configuring DHCP snooping](#)."

Enabling Option 82 handling

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable the server to handle Option 82.	dhcp server relay information enable	Optional. Enabled by default.

Specifying the threshold for sending trap messages

Configuration prerequisites

Before you perform the configuration, use the **snmp-agent target-host** command to specify the destination address of the trap messages. For more information about the command, see the *Network Management and Monitoring Command Reference*.

Configuration procedure

A DHCP server sends trap messages to the network management server when one of the following items reaches the specified threshold:

- The ratio of successfully allocated IP addresses to received DHCP requests
- The average IP address utilization of the address pool
- The maximum IP address utilization of the address pool

Trap messages help network administrators know the latest usage information of the DHCP server.

To specify the threshold for sending trap messages:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Specify the threshold for sending trap messages to the network management server.	dhcp server threshold { allocated-ip <i>threshold-value</i> average-ip-use <i>threshold-value</i> max-ip-use <i>threshold-value</i> }	Optional. Disabled by default.

Setting the DSCP value for DHCP packets

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the DSCP value for DHCP packets sent by the DHCP server.	dhcp dscp <i>dscp-value</i>	Optional. By default, the DSCP value is 56.

Displaying and maintaining the DHCP server

⚠ IMPORTANT:

A restart of the DHCP server or execution of the **reset dhcp server ip-in-use** command deletes all lease information. The DHCP server denies any DHCP request for lease extension, and the client must request an IP address again.

Task	Command	Remarks
Display information about IP address conflicts.	display dhcp server conflict { all ip <i>ip-address</i> } [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display information about lease expiration.	display dhcp server expired { all ip ip-address pool [<i>pool-name</i>] } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about assignable IP addresses.	display dhcp server free-ip [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IP addresses excluded from automatic allocation in the DHCP address pool.	display dhcp server forbidden-ip [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about bindings.	display dhcp server ip-in-use { all ip ip-address pool [<i>pool-name</i>] } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about DHCP server statistics.	display dhcp server statistics [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display tree organization information of address pool(s).	display dhcp server tree { all pool [<i>pool-name</i>] } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear information about IP address conflicts.	reset dhcp server conflict { all ip ip-address }	Available in user view
Clear information about dynamic bindings.	reset dhcp server ip-in-use { all ip ip-address pool [<i>pool-name</i>] }	Available in user view
Clear information about DHCP server statistics.	reset dhcp server statistics	Available in user view

DHCP server configuration examples

DHCP networking involves the following two types:

- The DHCP server and client are on the same subnet and exchange messages directly.
- The DHCP server and client are not on the same subnet and they communicate with each other via a DHCP relay agent.

The DHCP server configuration for the two types is the same.

Static IP address assignment configuration example

Network requirements

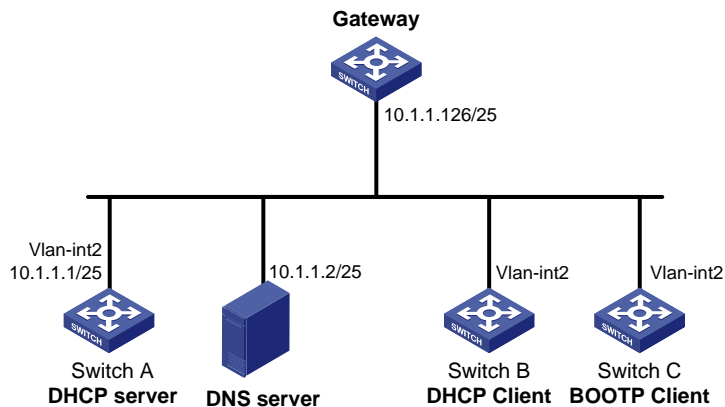
As shown in [Figure 29](#), Switch B (DHCP client) and Switch C (BOOTP client) obtain the static IP address, DNS server address, and gateway address from Switch A (DHCP server).

The client ID of VLAN-interface 2 on Switch B is:

3030-3066-2e65-3234-392e-3830-3530-2d56-6c61-6e2d-696e-7465-7266-6163-6532.

The MAC address of VLAN-interface 2 on Switch C is 000f-e249-8050.

Figure 29 Network diagram



Configuration procedure

1. Configure the IP address of VLAN-interface 2 on Switch A.

```
<SwitchA> system-view
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ip address 10.1.1.1 25
[SwitchA-Vlan-interface2] quit
```

2. Configure the DHCP server:

Enable DHCP.

```
[SwitchA] dhcp enable
```

Enable the DHCP server on VLAN-interface 2.

```
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] dhcp select server global-pool
[SwitchA-Vlan-interface2] quit
```

Create DHCP address pool 0, configure a static binding, DNS server and gateway in it.

```
[SwitchA] dhcp server ip-pool 0
[SwitchA-dhcp-pool-0] static-bind ip-address 10.1.1.5
[SwitchA-dhcp-pool-0] static-bind client-identifier
3030-3066-2e65-3234-392e-3830-3530-2d56-6c61-6e2d-696e-7465-7266-6163-6532
[SwitchA-dhcp-pool-0] dns-list 10.1.1.2
[SwitchA-dhcp-pool-0] gateway-list 10.1.1.126
[SwitchA-dhcp-pool-0] quit
```

Create DHCP address pool 1, configure a static binding, DNS server and gateway in it.

```
[SwitchA] dhcp server ip-pool 1
[SwitchA-dhcp-pool-1] static-bind ip-address 10.1.1.6
[SwitchA-dhcp-pool-1] static-bind mac-address 000f-e249-8050
[SwitchA-dhcp-pool-1] dns-list 10.1.1.2
[SwitchA-dhcp-pool-1] gateway-list 10.1.1.126
```

Verifying the configuration

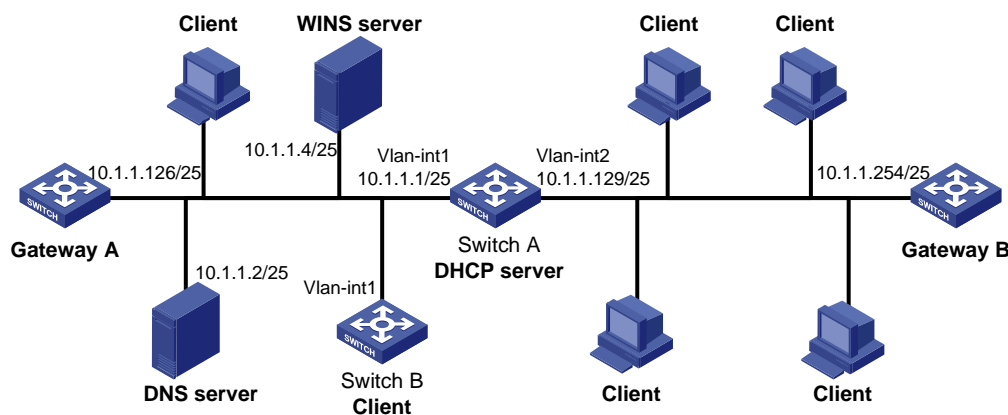
After the preceding configuration is complete, Switch B can obtain IP address 10.1.1.5 and other network parameters, and Switch C can obtain IP address 10.1.1.6 and other network parameters from Switch A. You can use the **display dhcp server ip-in-use** command on the DHCP server to view the IP addresses assigned to the clients.

Dynamic IP address assignment configuration example

Network requirements

- As shown in Figure 30, the DHCP server (Switch A) assigns IP addresses to clients in subnet 10.1.1.0/24, which is subnetted into 10.1.1.0/25 and 10.1.1.128/25.
- The IP addresses of VLAN-interfaces 1 and 2 on Switch A are 10.1.1.1/25 and 10.1.1.129/25 respectively.
- In address pool 10.1.1.0/25, configure the address lease duration as ten days and twelve hours, domain name suffix aabbcc.com, DNS server address 10.1.1.2/25, gateway 10.1.1.126/25, and WINS server 10.1.1.4/25.
- In address pool 10.1.1.128/25, configure the address lease duration as five days, domain name suffix aabbcc.com, DNS server address 10.1.1.2/25, and gateway address 10.1.1.254/25, and there is no WINS server address.
- The domain name and DNS server address on subnets 10.1.1.0/25 and 10.1.1.128/25 are the same. Therefore, the domain name suffix and DNS server address can be configured only for subnet 10.1.1.0/24. Subnet 10.1.1.128/25 can inherit the configuration of subnet 10.1.1.0/24.

Figure 30 Network diagram



Configuration procedure

1. Specify IP addresses for VLAN interfaces. (Details not shown.)
2. Configure the DHCP server:

Enable DHCP.

```
<SwitchA> system-view
[SwitchA] dhcp enable
```

Enable the DHCP server on VLAN-interface 1 and VLAN-interface 2.

```
[SwitchA] interface vlan-interface 1
[SwitchA-Vlan-interface1] dhcp select server global-pool
[SwitchA-Vlan-interface1] quit
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] dhcp select server global-pool
[SwitchA-Vlan-interface2] quit
```

Exclude IP addresses (addresses of the DNS server, WINS server and gateways).

```
[SwitchA] dhcp server forbidden-ip 10.1.1.2
[SwitchA] dhcp server forbidden-ip 10.1.1.4
```

```

[SwitchA] dhcp server forbidden-ip 10.1.1.126
[SwitchA] dhcp server forbidden-ip 10.1.1.254
# Configure DHCP address pool 0 (subnet, client domain name suffix, and DNS server address).
[SwitchA] dhcp server ip-pool 0
[SwitchA-dhcp-pool-0] network 10.1.1.0 mask 255.255.255.0
[SwitchA-dhcp-pool-0] domain-name aabbcc.com
[SwitchA-dhcp-pool-0] dns-list 10.1.1.2
[SwitchA-dhcp-pool-0] quit
# Configure DHCP address pool 1 (subnet, gateway, lease duration, and WINS server).
[SwitchA] dhcp server ip-pool 1
[SwitchA-dhcp-pool-1] network 10.1.1.0 mask 255.255.255.128
[SwitchA-dhcp-pool-1] gateway-list 10.1.1.126
[SwitchA-dhcp-pool-1] expired day 10 hour 12
[SwitchA-dhcp-pool-1] nbns-list 10.1.1.4
[SwitchA-dhcp-pool-1] quit
# Configure DHCP address pool 2 (subnet, gateway, and lease duration).
[SwitchA] dhcp server ip-pool 2
[SwitchA-dhcp-pool-2] network 10.1.1.128 mask 255.255.255.128
[SwitchA-dhcp-pool-2] expired day 5
[SwitchA-dhcp-pool-2] gateway-list 10.1.1.254

```

Verifying the configuration

After the preceding configuration is complete, clients on networks 10.1.1.0/25 and 10.1.1.128/25 can obtain IP addresses on the corresponding network and other network parameters from Switch A. You can use the **display dhcp server ip-in-use** command on the DHCP server to view the IP addresses assigned to the clients.

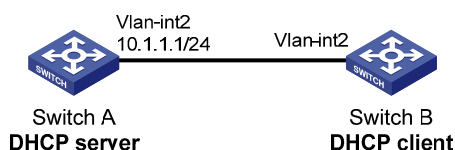
Self-defined option configuration example

Network requirements

As shown in [Figure 31](#), the DHCP client (Switch B) obtains an IP address and PXE server addresses from the DHCP server (Switch A). The IP address belongs to subnet 10.1.1.0/24. The PXE server addresses are 1.2.3.4 and 2.2.2.2.

The DHCP server assigns PXE server addresses to DHCP clients through Option 43, a self-defined option. The format of Option 43 and that of the PXE server address sub-option are shown in [Figure 19](#) and [Figure 21](#), respectively. The value of Option 43 configured on the DHCP server in this example is 80 0B 00 00 02 01 02 03 04 02 02 02 02. The number 80 is the value of the sub-option type. The number 0B is the value of the sub-option length. The numbers 00 00 are the value of the PXE server type. The number 02 indicates the number of servers. The numbers 01 02 03 04 02 02 02 02 indicate that the PXE server addresses are 1.2.3.4 and 2.2.2.2.

Figure 31 Network diagram



Configuration procedure

1. Specify IP addresses for the interfaces. (Details not shown.)
2. Configure the DHCP server:

```
# Enable DHCP.
<SwitchA> system-view
[SwitchA] dhcp enable

# Enable the DHCP server on VLAN-interface 2.
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] dhcp select server global-pool
[SwitchA-Vlan-interface2] quit

# Configure DHCP address pool 0.
[SwitchA] dhcp server ip-pool 0
[SwitchA-dhcp-pool-0] network 10.1.1.0 mask 255.255.255.0
[SwitchA-dhcp-pool-0] option 43 hex 80 0B 00 00 02 01 02 03 04 02 02 02
```

Verifying the configuration

After the preceding configuration is complete, Switch B can obtain its IP address on 10.1.1.0/24 and PXE server addresses from the Switch A. You can use the **display dhcp server ip-in-use** command on the DHCP server to view the IP addresses assigned to the clients.

Troubleshooting DHCP server configuration

Symptom

A client's IP address obtained from the DHCP server conflicts with another IP address.

Analysis

A host on the subnet may have the same IP address.

Solution

1. Disable the client's network adapter or disconnect the client's network cable. Ping the IP address of the client from another host to check whether there is a host using the same IP address.
2. If a ping response is received, the IP address has been manually configured on a host. Execute the **dhcp server forbidden-ip** command on the DHCP server to exclude the IP address from dynamic allocation.
3. Enable the network adapter or connect the network cable. For example, to release the IP address and obtain another one on a Windows XP client:
 - a. Run **cmd** to enter DOS window.
 - b. Type **ipconfig/release** to relinquish the IP address.
 - c. Type **ipconfig/renew** to obtain another IP address.

Configuring DHCP relay agent

The DHCP relay agent configuration is supported only on Layer 3 Ethernet ports, Layer 3 aggregate interfaces, and VLAN interfaces.

The term "interface" in the DHCP features refers to Layer 3 interfaces, including VLAN interfaces and route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see the *Layer 2—LAN Switching Configuration Guide*).

Only HP 5500 EI switches support Layer 3 Ethernet port configuration.

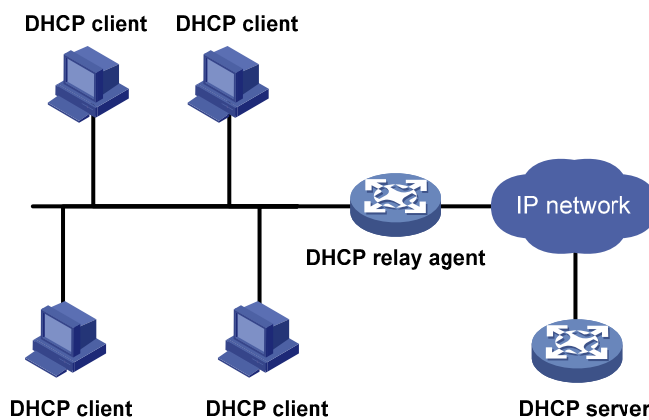
Overview

Via a relay agent, DHCP clients can communicate with a DHCP server on another subnet to obtain configuration parameters. DHCP clients on different subnets can contact the same DHCP server rather than having a DHCP server on each subnet. This centralizes management and reduces cost reduction.

An MCE device serving as the DHCP relay agent can forward DHCP packets not only between a DHCP server and clients on a public network, but also between a DHCP server and clients on a private network. Note that the IP address ranges of the public and private networks or those of private networks cannot overlap each other. For more information about MCE, see *Layer 3—IP Routing Configuration Guide*.

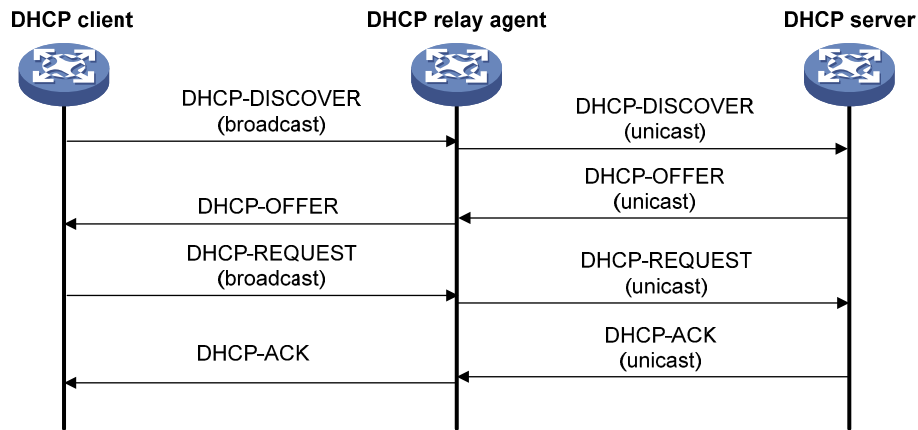
Fundamentals

Figure 32 DHCP relay agent application



The DHCP server and client interact with each other in the same way with or without a relay agent (see "[DHCP overview](#)").

Figure 33 DHCP relay agent work process



1. After receiving a DHCP-DISCOVER or DHCP-REQUEST broadcast message from a DHCP client, the DHCP relay agent fills the giaddr field of the message with its IP address and forwards the message to the designated DHCP server in unicast mode.
2. Based on the giaddr field, the DHCP server returns an IP address and other configuration parameters to the relay agent, and the relay agent conveys them to the client.

DHCP relay agent support for Option 82

Option 82 records location information of the DHCP client, letting the administrator locate the DHCP client for security control and accounting purposes. For more information, see "[DHCP overview](#)"

If the DHCP relay agent supports Option 82, it handles a client's request according to the contents defined in Option 82, if any. The handling strategies are described in [Table 3](#).

If a reply returned by the DHCP server contains Option 82, the DHCP relay agent removes the Option 82 before forwarding the reply to the client.

Table 3 Handling strategies of the DHCP relay agent

If a client's requesting message has...	Handling strategy	Padding format	The DHCP relay agent will...
	Drop	Random	Drop the message.
	Keep	Random	Forward the message without changing Option 82.
Option 82		normal	Forward the message after replacing the original Option 82 with the Option 82 padded in normal format.
	Replace	verbose	Forward the message after replacing the original Option 82 with the Option 82 padded in verbose format.
		user-defined	Forward the message after replacing the original Option 82 with the user-defined Option 82.
no Option 82	N/A	normal	Forward the message after adding the Option 82 padded in normal format.

If a client's requesting message has...	Handling strategy	Padding format	The DHCP relay agent will...
	N/A	verbose	Forward the message after adding the Option 82 padded in verbose format.
	N/A	user-defined	Forward the message after adding the user-defined Option 82.

DHCP relay agent configuration task list

Task	Remarks
Enabling DHCP	Required
Enabling the DHCP relay agent on an interface	Required
Correlating a DHCP server group with a relay agent interface	Required
Configuring the DHCP relay agent security functions	Optional
Enabling offline detection	Optional
Configuring the DHCP relay agent to release an IP address	Optional
Configuring the DHCP relay agent to support Option 82	Optional
Setting the DSCP value for DHCP packets	Optional

Enabling DHCP

Enable DHCP before performing other configurations related to the DHCP relay agent.

To enable DHCP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable DHCP.	dhcp enable	Disabled by default

Enabling the DHCP relay agent on an interface

With the DHCP relay agent enabled, an interface forwards incoming DHCP requests to a DHCP server for address allocation.

The IP address pool containing the IP address of the DHCP relay agent enabled interface must be configured on the DHCP server. Otherwise, the DHCP clients connected to the relay agent cannot obtain correct IP addresses.

To enable the DHCP relay agent on an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable the DHCP relay agent on the current interface.	dhcp select relay	With DHCP enabled, interfaces operate in the DHCP server mode.

Correlating a DHCP server group with a relay agent interface

To improve reliability, you can specify several DHCP servers as a group on the DHCP relay agent and correlate a relay agent interface with the server group. When the interface receives request messages from clients, the relay agent will forward them to all the DHCP servers of the group.

Configuration guidelines

Follow these guidelines when you correlate a DHCP server group with a relay agent interface:

- You can specify up to twenty DHCP server groups on the relay agent.
- By executing the **dhcp relay server-group** command repeatedly, you can specify up to eight DHCP server addresses for each DHCP server group.
- The IP addresses of DHCP servers and those of relay agent's interfaces that connect DHCP clients cannot be on the same subnet. Otherwise, the client cannot obtain an IP address.
- A DHCP server group can correlate with one or multiple DHCP relay agent interfaces, while a relay agent interface can only correlate with one DHCP server group. Using the **dhcp relay server-select** command repeatedly overwrites the previous configuration. However, if the specified DHCP server group does not exist, the interface still uses the previous correlation.
- The *group-id* argument in the **dhcp relay server-select** command is configured by using the **dhcp relay server-group** command.

Configuration procedure

To correlate a DHCP server group with a relay agent interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a DHCP server group and add a server into the group.	dhcp relay server-group <i>group-id</i> ip <i>ip-address</i>	Not created by default.
3. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Correlate the DHCP server group with the current interface.	dhcp relay server-select <i>group-id</i>	By default, no interface is correlated with any DHCP server group.

Configuring the DHCP relay agent security functions

Configuring address check

Address check can block illegal hosts from accessing external networks.

With this feature enabled, the DHCP relay agent can dynamically record clients' IP-to-MAC bindings after they obtain IP addresses through DHCP. This feature also supports static bindings. You can also configure static IP-to-MAC bindings on the DHCP relay agent, so users can access external networks using fixed IP addresses.

Upon receiving a packet from a host, the DHCP relay agent checks the source IP and MAC addresses in the packet against the recorded dynamic and static bindings. If no match is found, the DHCP relay agent does not learn the ARP entry of the host, and will not forward any reply to the host, so the host cannot access external networks via the DHCP relay agent.

Configuration guidelines

Follow these guidelines when you create a static binding and enable address check:

- The **dhcp relay address-check enable** command can be executed only on Layer 3 Ethernet ports and VLAN interfaces.
- Before enabling address check on an interface, you must enable the DHCP service, and enable the DHCP relay agent on the interface; otherwise, the address check configuration is ineffective.
- The **dhcp relay address-check enable** command only checks IP and MAC addresses but not interfaces.
- When using the **dhcp relay security static** command to bind an interface to a static binding entry, make sure that the interface is configured as a DHCP relay agent; otherwise, address entry conflicts may occur.

Configuration procedure

To create a static binding and enable address check:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a static binding.	dhcp relay security static <i>ip-address mac-address</i> [interface <i>interface-type interface-number</i>]	Optional. No static binding is created by default.
3. Enter interface view.	interface <i>interface-type interface-number</i>	N/A
4. Enable address check.	dhcp relay address-check enable	Disabled by default.

Configuring periodic refresh of dynamic client entries

A DHCP client unicasts a DHCP-RELEASE message to the DHCP server to release its IP address. The DHCP relay agent simply conveys the message to the DHCP server and does not remove the IP-to-MAC entry of the client.

When this feature is enabled, the DHCP relay agent uses the IP address of a client and the MAC address of the DHCP relay interface to send a DHCP-REQUEST message to the DHCP server at specified intervals.

- If the server returns a DHCP-ACK message or does not return any message within a specified interval, the DHCP relay agent ages out the entry.
- If the server returns a DHCP-NAK message, the relay agent keeps the entry.

To configure periodic refresh of dynamic client entries:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable periodic refresh of dynamic client entries.	dhcp relay security refresh enable	Optional. Enabled by default.
3. Configure the refresh interval.	dhcp relay security tracker { <i>interval</i> auto }	Optional. auto by default. (auto interval is calculated by the relay agent according to the number of client entries.)

Enabling unauthorized DHCP server detection

Unauthorized DHCP servers may assign wrong IP addresses to DHCP clients.

With unauthorized DHCP servers detection enabled, the DHCP relay agent checks whether a request contains Option 54 (Server Identifier Option). If yes, the DHCP relay agent records the IP address of each detected DHCP server that assigned an IP address to a requesting DHCP client in the option, and records the receiving interface. The administrator can use this information to check for unauthorized DHCP servers.

The relay agent logs a DHCP server only once.

To enable unauthorized DHCP server detection:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable unauthorized DHCP server detection.	dhcp relay server-detect	Disabled by default

Enabling DHCP starvation attack protection

A DHCP starvation attack occurs when an attacker constantly sends forged DHCP requests using different MAC addresses in the chaddr field to a DHCP server. This exhausts the IP address resources of the DHCP server so legitimate DHCP clients cannot obtain IP addresses. The DHCP server may also fail to work because of exhaustion of system resources.

- To relieve a DHCP starvation attack that uses DHCP packets encapsulated with different source MAC addresses, you can limit the number of ARP entries that a Layer 3 interface can learn or MAC addresses that a Layer 2 port can learn. You can also configure an interface that has learned the maximum MAC addresses to discard packets whose source MAC addresses are not in the MAC address table.

- To prevent a DHCP starvation attack that uses DHCP requests encapsulated with the same source MAC address, enable MAC address check on the DHCP relay agent. With this function enabled, the DHCP relay agent compares the chaddr field of a received DHCP request with the source MAC address field of the frame. If they are the same, the DHCP relay agent decides this request as valid and forwards it to the DHCP server; if not, it discards the DHCP request.

To enable MAC address check:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable MAC address check.	dhcp relay check mac-address	Disabled by default

NOTE:

DHCP relay agents change the source MAC addresses when forwarding DHCP packets. Therefore, you can enable MAC address check only on a DHCP relay agent directly connected to DHCP clients. Otherwise, valid DHCP packets may be discarded and clients cannot obtain IP addresses.

Enabling offline detection

The DHCP relay agent checks whether a user is online by learning the ARP entry. When an ARP entry is aged out, the corresponding client is considered to be offline.

With this function enabled on an interface, the DHCP relay agent removes a client's IP-to-MAC entry when it is aged out, and sends a DHCP-RELEASE message to the DHCP server to release the IP address of the client. Removing an ARP entry manually does not remove the corresponding client's IP-to-MAC binding. When the client goes offline, use the **undo dhcp relay security** command to remove the IP-to-MAC binding manually.

To enable offline detection:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable offline detection.	dhcp relay client-detect enable	Disabled by default

Configuring the DHCP relay agent to release an IP address

You can configure the relay agent to release a client's IP address. The relay agent sends a DHCP-RELEASE message that contains the IP address. Upon receiving the DHCP-RELEASE message, the DHCP server releases the IP address; meanwhile, the client entry is removed from the DHCP relay agent. Dynamic client entries can be generated after you enable address check or IP source guard on the DHCP relay agent. For more information about IP source guard, see the *Security Configuration Guide*.

To configure the DHCP relay agent to send DHCP-RELEASE messages:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the DHCP relay agent to release an IP address.	dhcp relay release ip <i>client-ip</i>	The IP address must be in a dynamic client entry.

Configuring the DHCP relay agent to support Option 82

Configuration prerequisites

Before you perform this configuration, complete the following tasks:

- Enable DHCP.
- Enable the DHCP relay agent on the specified interface.
- Correlate a DHCP server group with relay agent interfaces.

Configuration guidelines

- To support Option 82, perform related configuration on both the DHCP server and relay agent. See "[Configuring DHCP server](#)" for DHCP server configuration of this kind.
- If the handling strategy of the DHCP relay agent is configured as **replace**, you must configure a padding format for Option 82. If the handling strategy is **keep** or **drop**, you need not configure any padding format.
- If sub-option 1 (node identifier) of Option 82 is padded with the device name (sysname) of a node, the device name must contain no spaces. Otherwise, the DHCP relay agent will drop the message.

Configuration procedure

To configure the DHCP relay agent to support Option 82:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable the relay agent to support Option 82.	dhcp relay information enable	Disabled by default.
4. Configure the handling strategy for requesting messages containing Option 82.	dhcp relay information strategy { drop keep replace }	Optional. replace by default.

Step	Command	Remarks
5. Configure non-user-defined Option 82.	<ul style="list-style-type: none"> Configure the padding format for Option 82: dhcp relay information format { normal verbose [node-identifier { mac sysname user-defined node-identifier }] } Configure the code type for the circuit ID sub-option: dhcp relay information circuit-id format-type { ascii hex } Configure the code type for the remote ID sub-option: dhcp relay information remote-id format-type { ascii hex } 	<p>Optional.</p> <p>By default,</p> <ul style="list-style-type: none"> The padding format for Option 82 is normal. The code type for the circuit ID sub-option depends on the padding format of Option 82. Each field has its own code type. The code type for the remote ID sub-option is hex. <p>The code type configurations for the circuit ID sub-option and remote ID sub-option apply to non-user-defined Option 82 only.</p>
6. Configure user-defined Option 82.	<ul style="list-style-type: none"> Configure the padding content for the circuit ID sub-option: dhcp relay information circuit-id string <i>circuit-id</i> Configure the padding content for the remote ID sub-option: dhcp relay information remote-id string { <i>remote-id</i> sysname } 	<p>Optional.</p> <p>By default, the padding content depends on the padding format of Option 82.</p>

Setting the DSCP value for DHCP packets

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the DSCP value for DHCP packets sent by the DHCP relay agent.	dhcp dscp <i>dscp-value</i>	<p>Optional.</p> <p>By default, the DSCP value is 56.</p>

Displaying and maintaining the DHCP relay agent

Task	Command	Remarks
Display information about DHCP server groups correlated to a specified interface or all interfaces.	display dhcp relay { all interface <i>interface-type interface-number</i> } [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display Option 82 configuration information on the DHCP relay agent.	display dhcp relay information { all interface <i>interface-type interface-number</i> } [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display information about bindings of DHCP relay agents.	display dhcp relay security [<i>ip-address</i> dynamic static] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view

Task	Command	Remarks
Display statistics about bindings of DHCP relay agents.	display dhcp relay security statistics [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about the refreshing interval for entries of dynamic IP-to-MAC bindings.	display dhcp relay security tracker [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about the configuration of a specified DHCP server group or all DHCP server groups.	display dhcp relay server-group { <i>group-id</i> all } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display packet statistics on relay agent.	display dhcp relay statistics [server-group { <i>group-id</i> all }] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear packet statistics from relay agent.	reset dhcp relay statistics [server-group <i>group-id</i>]	Available in user view

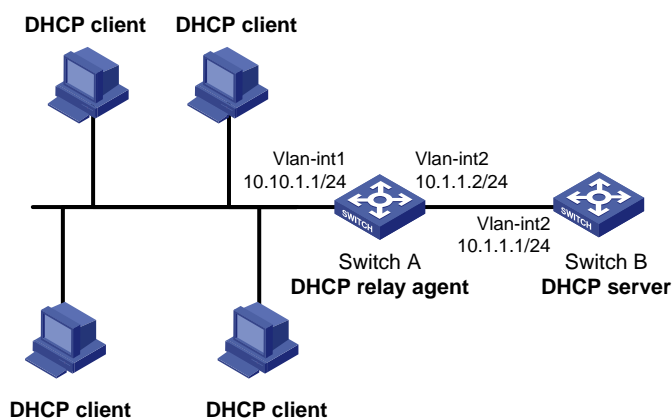
DHCP relay agent configuration examples

DHCP relay agent configuration example

Network requirements

As shown in [Figure 34](#), DHCP clients reside on network 10.10.1.0/24. The IP address of the DHCP server is 10.1.1.1/24. Because the DHCP clients reside on a different network than the DHCP server, a DHCP relay agent is deployed to forward messages between DHCP clients and the DHCP server. VLAN-interface 1 on the DHCP relay agent (Switch A) connects to the network where DHCP clients reside. The IP address of VLAN-interface 1 is 10.10.1.1/24 and the IP address of VLAN-interface 2 is 10.1.1.2/24.

Figure 34 Network diagram



Configuration procedure

The DHCP relay agent and server are on different subnets, so configure a static route or dynamic routing protocol to make them reachable to each other.

Configurations on the DHCP server are also required to guarantee the client-server communication via the DHCP relay agent. For DHCP server configuration information, see "[Configuring DHCP server.](#)"

Specify IP addresses for the interfaces. (Details not shown.)

Enable DHCP.

```
<SwitchA> system-view
```

```
[SwitchA] dhcp enable
```

Add DHCP server 10.1.1.1 into DHCP server group 1.

```
[SwitchA] dhcp relay server-group 1 ip 10.1.1.1
```

Enable the DHCP relay agent on VLAN-interface 1.

```
[SwitchA] interface vlan-interface 1
```

```
[SwitchA-Vlan-interface1] dhcp select relay
```

Correlate VLAN-interface 1 to DHCP server group 1.

```
[SwitchA-Vlan-interface1] dhcp relay server-select 1
```

After the preceding configuration is complete, DHCP clients can obtain IP addresses and other network parameters through the DHCP relay agent from the DHCP server. You can use the **display dhcp relay statistics** command to view statistics of DHCP packets forwarded by DHCP relay agents. After you enable address check of the DHCP relay agents with the **dhcp relay address-check enable** command, use the **display dhcp relay security** command to view bindings of DHCP relay agents

DHCP relay agent Option 82 support configuration example

Network requirements

- As shown in [Figure 34](#), enable Option 82 on the DHCP relay agent (Switch A).
- Configure the handling strategy for DHCP requests containing Option 82 as **replace**.
- Configure the padding content for the circuit ID sub-option as **company001** and for the remote ID sub-option as **device001**.
- Switch A forwards DHCP requests to the DHCP server (Switch B) after replacing Option 82 in the requests, so that the DHCP clients can obtain IP addresses.

Configuration procedure

Configurations on the DHCP server are also required to make the Option 82 configurations function normally.

Specify IP addresses for the interfaces. (Details not shown.)

Enable DHCP.

```
<SwitchA> system-view
```

```
[SwitchA] dhcp enable
```

Add DHCP server 10.1.1.1 into DHCP server group 1.

```
[SwitchA] dhcp relay server-group 1 ip 10.1.1.1
```

Enable the DHCP relay agent on VLAN-interface 1.

```
[SwitchA] interface vlan-interface 1
```

```
[SwitchA-Vlan-interface1] dhcp select relay
```

Correlate VLAN-interface 1 to DHCP server group 1.

```
[SwitchA-Vlan-interface1] dhcp relay server-select 1
```

Enable the DHCP relay agent to support Option 82, and perform Option 82-related configurations.

```
[SwitchA-Vlan-interface1] dhcp relay information enable
[SwitchA-Vlan-interface1] dhcp relay information strategy replace
[SwitchA-Vlan-interface1] dhcp relay information circuit-id string company001
[SwitchA-Vlan-interface1] dhcp relay information remote-id string device001
```

Troubleshooting DHCP relay agent configuration

Symptom

DHCP clients cannot obtain any configuration parameters via the DHCP relay agent.

Analysis

Problems may occur with the DHCP relay agent or server configuration.

Solution

To locate the problem, enable debugging and execute the **display** command on the DHCP relay agent to view the debugging information and interface state information.

Verify that:

- The DHCP is enabled on the DHCP server and relay agent.
- The address pool on the same subnet where DHCP clients reside is available on the DHCP server.
- The DHCP server and DHCP relay agent are reachable to each other.
- The relay agent interface connected to DHCP clients is correlated with a correct DHCP server group and the IP addresses of the group members are correct.

Configuring DHCP client

With DHCP client enabled, an interface uses DHCP to obtain configuration parameters such as an IP address from the DHCP server.

Configuration restrictions

- The DHCP client configuration is supported only on Layer 3 Ethernet ports, Layer 3 aggregate interfaces, and VLAN interfaces.
- When multiple VLAN interfaces with the same MAC address use DHCP for IP address acquisition via a relay agent, the DHCP server cannot be a Windows Server 2000 or Windows Server 2003.
- You cannot configure an interface of an aggregation group as a DHCP client.
- Only HP 5500 EI switches support Layer 3 Ethernet port configuration.

Enabling the DHCP client on an interface

Follow these guidelines when you enable the DHCP client on an interface:

- An interface can be configured to acquire an IP address in multiple ways. The latest configuration overwrites the previous one.
- Secondary IP addresses cannot be configured on an interface that is enabled with the DHCP client.
- If the IP address that interface A obtains from the DHCP server is on the same network segment as the IP address of interface B, interface A neither uses the IP address nor requests any IP address from the DHCP server unless you do the following: Delete the IP address of interface B and bring up interface A again by first executing the **shutdown** command and then the **undo shutdown** command, or, re-enable the DHCP client on interface A by executing the **undo ip address dhcp-alloc** command and then the **ip address dhcp-alloc** command.

To enable the DHCP client on an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type interface-number</i>	N/A
3. Enable the DHCP client on the interface.	ip address dhcp-alloc [client-identifier mac <i>interface-type interface-number</i>]	Disabled by default

Setting the DSCP value for DHCP packets

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the DSCP value for DHCP packets sent by the DHCP client.	dhcp client dscp <i>dscp-value</i>	Optional. By default, the DSCP value is 56.

Displaying and maintaining the DHCP client

Task	Command	Remarks
Display specified configuration information.	display dhcp client [verbose] [interface <i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

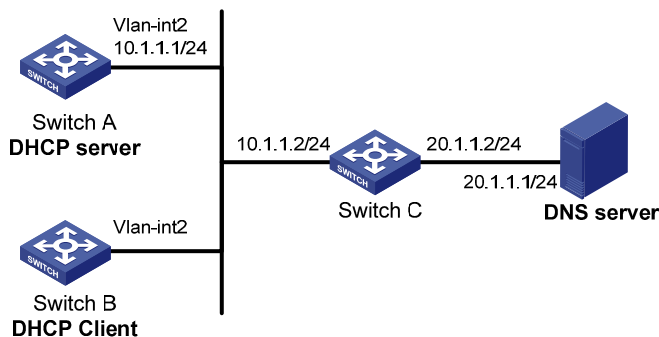
DHCP client configuration example

Network requirements

As shown in [Figure 35](#), on a LAN, Switch B contacts the DHCP server via VLAN-interface 2 to obtain an IP address, DNS server address, and static route information. The IP address resides on network 10.1.1.0/24. The DNS server address is 20.1.1.1. The next hop of the static route to network 20.1.1.0/24 is 10.1.1.2.

The DHCP server uses Option 121 to assign static route information to DHCP clients. The destination descriptor field comprises two parts, subnet mask length and destination network address. In this example, the value of the destination descriptor field takes 18 14 01 01, a hexadecimal number indicating that the subnet mask length is 24 and destination network address is 20.1.1.0. The value of the next hop address field takes 0A 01 01 02, a hexadecimal number indicating that the next hop is 10.1.1.2.

Figure 35 Network diagram



Configuration procedure

1. Configure Switch A:

Specify the IP address of VLAN-interface 2.

```
<SwitchA> system-view
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ip address 10.1.1.1 24
[SwitchA-Vlan-interface2] quit
```

Enable the DHCP service.

```
[SwitchA] dhcp enable
```

Exclude an IP address from automatic allocation.

```
[SwitchA] dhcp server forbidden-ip 10.1.1.2
```

Configure DHCP address pool 0 and specify the subnet, lease duration, DNS server address, and a static route to subnet 20.1.1.0/24.

```
[SwitchA] dhcp server ip-pool 0
[SwitchA-dhcp-pool-0] network 10.1.1.0 mask 255.255.255.0
[SwitchA-dhcp-pool-0] expired day 10
[SwitchA-dhcp-pool-0] dns-list 20.1.1.1
[SwitchA-dhcp-pool-0] option 121 hex 18 14 01 01 0A 01 01 02
```

2. Configure Switch B:

Enable the DHCP client on VLAN-interface 2.

```
<SwitchB> system-view
[SwitchB] interface vlan-interface 2
[SwitchB-Vlan-interface2] ip address dhcp-alloc
```

Verifying the configuration

Use the **display dhcp client** command to view the IP address and other network parameters assigned to Switch B.

```
[SwitchB-Vlan-interface2] display dhcp client verbose
Vlan-interface2 DHCP client information:
Current machine state: BOUND
Allocated IP: 10.1.1.3 255.255.255.0
Allocated lease: 864000 seconds, T1: 432000 seconds, T2: 756000 seconds
Lease from 2009.02.20 11:06:35 to 2009.03.02 11:06:35
DHCP server: 10.1.1.1
Transaction ID: 0x410090f0
Classless static route:
  Destination: 20.1.1.0, Mask: 255.255.255.0, NextHop: 10.1.1.2
DNS server: 20.1.1.1
Client ID: 3030-3066-2e65-3230-
          302e-3030-3032-2d45-
          7468-6572-6e65-7430-
          2f30
T1 will timeout in 4 days 23 hours 59 minutes 50 seconds.
```

Use the **display ip routing-table** command to view the route information on Switch B. A static route to network 20.1.1.0/24 is added to the routing table.

```
[SwitchB-Vlan-interface2] display ip routing-table
Routing Tables: Public
          Destinations : 5          Routes : 5

Destination/Mask    Proto  Pre  Cost           NextHop         Interface
-----
10.1.1.0/24         Direct  0    0              10.1.1.3         Vlan2
10.1.1.3/32         Direct  0    0              127.0.0.1        InLoop0
20.1.1.0/24         Static  70   0              10.1.1.2         Vlan2
127.0.0.0/8         Direct  0    0              127.0.0.1        InLoop0
127.0.0.1/32        Direct  0    0              127.0.0.1        InLoop0
```

Configuring DHCP snooping

The DHCP snooping-enabled device must be either between the DHCP client and relay agent, or between the DHCP client and server. It does not work if it is between the DHCP relay agent and DHCP server.

DHCP snooping functions

DHCP snooping can:

1. Ensure that DHCP clients obtain IP addresses from authorized DHCP servers.
2. Record IP-to-MAC mappings of DHCP clients.

Ensuring that DHCP clients obtain IP addresses from authorized DHCP servers

With DHCP snooping, the ports of a switch can be configured as trusted or untrusted to make sure that clients obtain IP addresses only from authorized DHCP servers.

- **Trusted**—A trusted port forwards DHCP messages normally to ensure the clients get IP addresses from an authorized DHCP server.
- **Untrusted**—An untrusted port discards received DHCP-ACK and DHCP-OFFER messages to avoid IP address allocation from any unauthorized server.

Configure ports that connect to authorized DHCP servers or other DHCP snooping devices as trusted, and configure other ports as untrusted.

Recording IP-to-MAC mappings of DHCP clients

DHCP snooping reads DHCP-REQUEST messages and DHCP-ACK messages from trusted ports to record DHCP snooping entries. A DHCP snooping entry includes the MAC and IP addresses of the client, the port that connects to the DHCP client, and the VLAN of the port. Using DHCP snooping entries, DHCP snooping can implement the following functions:

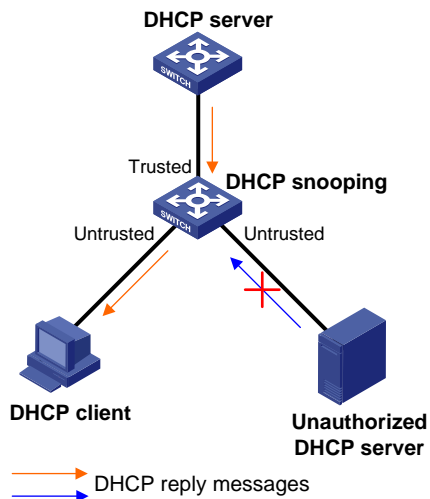
- **ARP detection**—Whether ARP packets are sent from an authorized client is determined based on DHCP snooping entries. This feature prevents ARP attacks from unauthorized clients. For more information, see the *Security Configuration Guide*.
- **IP source guard**—IP source guard uses dynamic binding entries generated by DHCP snooping to filter packets on a per-port basis. This prevents unauthorized packets from traveling through. For more information, see the *Security Configuration Guide*.
- **VLAN mapping**—The device replaces service provider VLANs (SVLANs) in packets with customer VLANs (CVLANs) by searching corresponding DHCP snooping entries for DHCP client information including IP addresses, MAC addresses, and CVLANs, before sending the packets to clients. For more information, see *Layer 2—LAN Switching Configuration Guide*.

Application environment of trusted ports

Configuring a trusted port connected to a DHCP server

As shown in Figure 36, the DHCP snooping device port that is connected to an authorized DHCP server should be configured as a trusted port. The trusted port forwards reply messages from the authorized DHCP server to the client, but the untrusted port does not forward reply messages from the unauthorized DHCP server. This ensures that the DHCP client obtains an IP address from the authorized DHCP server.

Figure 36 Configuring trusted and untrusted ports



Configuring trusted ports in a cascaded network

In a cascaded network involving multiple DHCP snooping devices, the ports connected to other DHCP snooping devices should be configured as trusted ports.

To save system resources, you can disable the trusted ports, which are indirectly connected to DHCP clients, from recording client IP-to-MAC bindings upon receiving DHCP requests.

Figure 37 Configuring trusted ports in a cascaded network

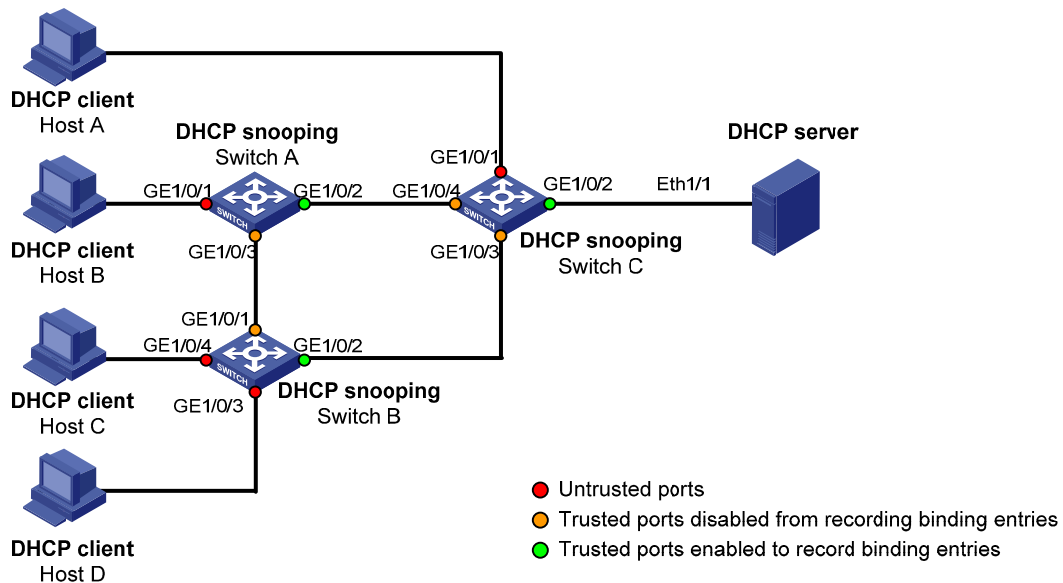


Table 4 Roles of ports

Device	Untrusted port	Trusted port disabled from recording binding entries	Trusted port enabled to record binding entries
Switch A	GigabitEthernet 1/0/1	GigabitEthernet 1/0/3	GigabitEthernet 1/0/2
Switch B	GigabitEthernet 1/0/3 and GigabitEthernet 1/0/4	GigabitEthernet 1/0/1	GigabitEthernet 1/0/2
Switch C	GigabitEthernet 1/0/1	GigabitEthernet 1/0/3 and GigabitEthernet 1/0/4	GigabitEthernet 1/0/2

DHCP snooping support for Option 82

Option 82 records the location information of the DHCP client so the administrator can locate the DHCP client for security control and accounting purposes. For more information, see "[Configuring DHCP relay agent.](#)"

If DHCP snooping supports Option 82, it handles a client's request according to the contents defined in Option 82, if any. The handling strategies are described in [Table 5](#).

If a reply returned by the DHCP server contains Option 82, the DHCP snooping device removes the Option 82 before forwarding the reply to the client. If the reply contains no Option 82, the DHCP snooping device forwards it directly.

Table 5 Handling strategies of DHCP snooping

If a client's requesting message has...	Handling strategy	Padding format	The DHCP snooping device will...
Option 82	Drop	N/A	Drop the message.
	Keep	Random	Forward the message without changing Option 82.

If a client's requesting message has...	Handling strategy	Padding format	The DHCP snooping device will...	
	Replace	normal	Forward the message after replacing the original Option 82 with the Option 82 padded in normal format.	
		verbose	Forward the message after replacing the original Option 82 with the Option 82 padded in verbose format.	
		user-defined	Forward the message after replacing the original Option 82 with the user-defined Option 82.	
	Append	normal	Forward the message without changing Option 82.	
		verbose	Forward the message without changing Option 82.	
		private	Forward the message after adding sub-option 9 to option 82 or adding content to sub-option 9 that option 82 contains.	
		standard	Forward the message without changing Option 82.	
		user-defined	Forward the message without changing Option 82.	
	no Option 82	N/A	normal	Forward the message after adding the Option 82 padded in normal format.
		N/A	private	Forward the message after adding the Option 82 padded in private format.
N/A		standard	Forward the message after adding the Option 82 padded in standard format.	
N/A		verbose	Forward the message after adding the Option 82 padded in verbose format.	
N/A		user-defined	Forward the message after adding the user-defined Option 82.	

The handling strategy and padding format for Option 82 on the DHCP snooping device are the same as those on the relay agent.

DHCP snooping configuration task list

Task	Remarks
Configuring DHCP snooping basic functions	Required
Configuring DHCP snooping to support Option 82	Optional
Configuring DHCP snooping entries backup	Optional
Enabling DHCP starvation attack protection	Optional

Task	Remarks
Enabling DHCP-REQUEST message attack protection	Optional
Configuring DHCP packet rate limit	Optional

Configuring DHCP snooping basic functions

Configuration guidelines

Follow these guidelines when configure DHCP snooping basic functions:

- You must specify the ports connected to the authorized DHCP servers as trusted to make sure that DHCP clients can obtain valid IP addresses. The trusted port and the port connected to the DHCP client must be in the same VLAN.
- You can specify Layer 2 Ethernet ports and Layer 2 aggregate interfaces as trusted ports. For more information about aggregate interfaces, see the *Layer 2—LAN Switching Configuration Guide*.
- If a Layer 2 Ethernet port is added to an aggregation group, the DHCP snooping configuration of the interface will not take effect. After the interface quits the aggregation group, the configuration will be effective.
- DHCP snooping can work with basic QinQ or flexible QinQ. When receiving a packet without any VLAN tag from the DHCP client to the DHCP server, the DHCP snooping device adds a VLAN tag to the packet. If the packet has one VLAN tag, the device adds another VLAN tag to the packet and records the two VLAN tags in a DHCP snooping entry. The newly added VLAN tag is the outer tag. If the packet has two VLAN tags, the device directly forwards the packet to the DHCP server without adding any tag.
- If you need to add a new VLAN tag and meanwhile modify the original VLAN tag for the packet, DHCP snooping cannot work with flexible QinQ.

Configuration procedure

To configure DHCP snooping basic functions:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable DHCP snooping.	dhcp-snooping	Disabled by default.
3. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	The interface connects to the DHCP server.
4. Specify the port as a trusted port that records the IP-to-MAC bindings of clients.	dhcp-snooping trust	After DHCP snooping is enabled, a port is an untrusted port by default
5. Return to system view.	quit	N/A
6. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	The interface indirectly connects to the DHCP client.
7. Specify the port as a trusted port that does not record the IP-to-MAC bindings of clients.	dhcp-snooping trust no-user-binding	Optional. After DHCP snooping is enabled, a port is an untrusted port by default.

Configuring DHCP snooping to support Option 82

Configuration guidelines

Follow these guidelines when configure DHCP snooping to support Option 82:

- You can only enable DHCP snooping to support Option 82 on Layer 2 Ethernet ports, and Layer 2 aggregate interfaces.
- If a Layer 2 Ethernet port is added to an aggregation group, enabling DHCP snooping to support Option 82 on the interface will not take effect. After the interface quits the aggregation group, the configuration will be effective.
- Option 82 support requires configuration on both the DHCP server and the device enabled with DHCP snooping. See "[Configuring DHCP server](#)" for DHCP server configuration of this kind.
- If the handling strategy of the DHCP-snooping-enabled device is configured as **replace**, you need to configure a padding format for Option 82. If the handling strategy is **keep** or **drop**, you need not configure any padding format.
- If the Option 82 is padded with the device name, the device name must contain no spaces. Otherwise, the DHCP-snooping device will drop the message. You can use the **sysname** command to specify the device name. For more information about this command, see the *Fundamentals Command Reference*.
- If DHCP snooping and QinQ work together or the DHCP snooping device receives a DHCP packet with two VLAN tags, and the normal or verbose padding format is adopted for Option 82, DHCP snooping fills the VLAN ID field of sub-option 1 with outer VLAN tag.inter VLAN tag. For example, if the outer VLAN tag is 10 (a in hexadecimal) and the inner VLAN tag is 20 (14 in hexadecimal), the VLAN ID is 000a.0014.

Configuration procedure

To configure DHCP snooping to support Option 82:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable DHCP snooping to support Option 82.	dhcp-snooping information enable	Disabled by default.
4. Configure the handling strategy for requests containing Option 82.	dhcp-snooping information strategy { append drop keep replace }	Optional. replace by default.

Step	Command	Remarks
5.	<p>Configure Option 82 in the non-user-defined padding format.</p> <ul style="list-style-type: none"> Configure the padding format for Option 82: dhcp-snooping information format { normal private <1-1> standard verbose [node-identifier { mac sysname user-defined node-identifier }] } Configure the code type for the circuit ID sub-option: dhcp-snooping information circuit-id format-type { ascii hex } Configure the code type for the remote ID sub-option: dhcp-snooping information remote-id format-type { ascii hex } Enable sub-option 9: dhcp-snooping information [vlan <i>vlan-id</i>] sub-option <i>sub-option-code</i> 	<p>Optional.</p> <p>By default,</p> <ul style="list-style-type: none"> The padding format for Option 82 is normal. The code type for the circuit ID sub-option depends on the padding format of Option 82. Each field has its own code type. The code type for the remote ID sub-option is hex. Sub-option 9 is not enabled <p>Hex configuration applies to private padding format only.</p> <p>The code type configuration for the circuit ID sub-option and remote ID sub-option apply to non-user-defined Option 82 only.</p> <p>For sub-option 9, when append strategy is adopted, the sysname and the primary IP address of the Loopback0 interface are padded. When some other strategy is adopted, only the sysname is padded.</p>
6.	<p>Configure user-defined Option 82.</p> <ul style="list-style-type: none"> Configure the padding content for the circuit ID sub-option: dhcp-snooping information [vlan <i>vlan-id</i>] circuit-id string <i>circuit-id</i> Configure the padding content for the remote ID sub-option: dhcp-snooping information [vlan <i>vlan-id</i>] remote-id string { <i>remote-id</i> sysname } Configure the padding content for the sub-option 9: dhcp-snooping information [vlan <i>vlan-id</i>] sub-option <i>sub-option-code</i> [string <i>user-string</i>&<1-8>] 	<p>Optional.</p> <p>By default,</p> <ul style="list-style-type: none"> The padding content for the circuit ID sub-option depends on the padding format of Option 82. The padding content for the remote ID sub-option depends on the padding format of Option 82. Sub-option 9 is not padded.

Configuring DHCP snooping entries backup

DHCP snooping entries cannot survive a reboot. If the DHCP snooping device is rebooted, security modules (such as IP source guard) that use DHCP snooping entries to authenticate users will reject requests from clients until new entries are learned.

The DHCP snooping entries backup feature enables you to store DHCP snooping entries in a file. When the DHCP snooping device reboots, it reads DHCP snooping entries from this file.

After DHCP snooping is disabled with the **undo dhcp-snooping** command, the device will delete all DHCP snooping entries, including those stored in the file.

To configure DHCP snooping entries backup:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Specify the name of the file for storing DHCP snooping entries.	dhcp-snooping binding database filename <i>filename</i>	Not specified by default. DHCP snooping entries are stored immediately after this command is used and then updated at the interval set by the dhcp-snooping binding database update interval command.
3. Back up DHCP snooping entries to the file.	dhcp-snooping binding database update now	Optional. DHCP snooping entries will be stored to the file each time this command is used.
4. Set the interval at which the DHCP snooping entry file is refreshed.	dhcp-snooping binding database update interval <i>minutes</i>	Optional. By default, the file is not refreshed periodically.

Enabling DHCP starvation attack protection

A DHCP starvation attack occurs when an attacker constantly sends forged DHCP requests using different MAC addresses in the chaddr field to a DHCP server. This exhausts the IP address resources of the DHCP server so legitimate DHCP clients cannot obtain IP addresses. The DHCP server may also fail to work because of exhaustion of system resources. You can protect against starvation attacks in the following ways:

- To relieve a DHCP starvation attack that uses DHCP packets encapsulated with different source MAC addresses, you can limit the number of MAC addresses that a Layer 2 port can learn.
- To prevent a DHCP starvation attack that uses DHCP requests encapsulated with the same source MAC address, enable MAC address check on the DHCP snooping device. With this function enabled, the DHCP snooping device compares the chaddr field of a received DHCP request with the source MAC address field of the frame. If they are the same, the request is considered valid and forwarded to the DHCP server; if not, the request is discarded.

Enable MAC address check only on Layer 2 Ethernet ports and Layer 2 aggregate interfaces.

To enable MAC address check:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable MAC address check.	dhcp-snooping check mac-address	Disabled by default

Enabling DHCP-REQUEST message attack protection

Attackers may forge DHCP-REQUEST messages to renew the IP address leases for legitimate DHCP clients that no longer need the IP addresses. These forged messages keep a victim DHCP server renewing the leases of IP addresses instead of releasing the IP addresses. This wastes IP address resources.

To prevent such attacks, you can enable DHCP-REQUEST message check on DHCP snooping devices. With this feature enabled, upon receiving a DHCP-REQUEST message, a DHCP snooping device looks up local DHCP snooping entries for the corresponding entry of the message. If an entry is found, the DHCP snooping device compares the entry with the message information. If they are consistent, the DHCP-REQUEST message is considered a valid lease renewal request and forwarded to the DHCP server. If they are not consistent, the message is considered a forged lease renewal request and discarded. If no corresponding entry is found, the message is considered valid and forwarded to the DHCP server.

Enable DHCP-REQUEST message check only on Layer 2 Ethernet ports, and Layer 2 aggregate interfaces.

To enable DHCP-REQUEST message check:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable DHCP-REQUEST message check.	dhcp-snooping check request-message	Disabled by default

Configuring DHCP packet rate limit

Configuration guidelines

- You can configure DHCP packet rate limit only on Layer 2 Ethernet ports and Layer 2 aggregate interfaces.
- If a Layer 2 Ethernet port belongs to an aggregation group, it uses the DHCP packet maximum rate configured on the corresponding Layer 2 aggregate interface.
- To identify DHCP packets from unauthorized DHCP servers, DHCP snooping delivers all incoming DHCP packets to the CPU. If a malicious user sends a large number of DHCP requests to the DHCP snooping device, the CPU of the device will be overloaded, and the device may even crash. To solve this problem, you can configure DHCP packet rate limit on relevant interfaces.

Configuration procedure

To configure DHCP packet rate limit:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet port view or Layer 2 aggregate interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A

Step	Command	Remarks
3. Configure the maximum rate of incoming DHCP packets.	dhcp-snooping rate-limit <i>rate</i>	Not configured by default

Displaying and maintaining DHCP snooping

Task	Command	Remarks
Display DHCP snooping entries.	display dhcp-snooping [<i>ip ip-address</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display Option 82 configuration information on the DHCP snooping device.	display dhcp-snooping information { all <i>interface interface-type</i> <i>interface-number</i> } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display DHCP packet statistics on the DHCP snooping device.	display dhcp-snooping packet statistics [<i>slot slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about trusted ports.	display dhcp-snooping trust [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the DHCP snooping entry file information.	display dhcp-snooping binding database [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear DHCP snooping entries.	reset dhcp-snooping { all <i>ip ip-address</i> }	Available in user view
Clear DHCP packet statistics on the DHCP snooping device.	reset dhcp-snooping packet statistics [<i>slot slot-number</i>]	Available in user view

DHCP snooping configuration examples

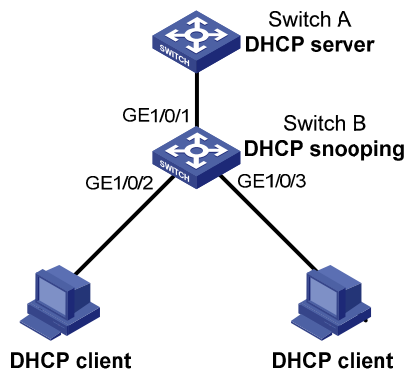
DHCP snooping configuration example

Network requirements

As shown in [Figure 38](#), Switch B is connected to a DHCP server through GigabitEthernet 1/0/1, and to two DHCP clients through GigabitEthernet 1/0/2 and GigabitEthernet 1/0/3. GigabitEthernet 1/0/1 forwards DHCP server responses while the other two do not.

Switch B records clients' IP-to-MAC address bindings in DHCP-REQUEST messages and DHCP-ACK messages received from trusted ports.

Figure 38 Network diagram



Configuration procedure

```
# Enable DHCP snooping.
<SwitchB> system-view
[SwitchB] dhcp-snooping

# Specify GigabitEthernet 1/0/1 as trusted.
[SwitchB] interface GigabitEthernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] dhcp-snooping trust
[SwitchB-GigabitEthernet1/0/1] quit
```

DHCP snooping Option 82 support configuration example

Network requirements

As shown in Figure 38, enable DHCP snooping and Option 82 support on Switch B.

- Configure the handling strategy for DHCP requests containing Option 82 as **replace**.
- On GigabitEthernet 1/0/2, configure the padding content for the circuit ID sub-option as **company001** and for the remote ID sub-option as **device001**.
- On GigabitEthernet 1/0/3, configure the padding format as **verbose**, access node identifier as **sysname**, and code type as **ascii** for Option 82.
- Switch B forwards DHCP requests to the DHCP server (Switch A) after replacing Option 82 in the requests, so that the DHCP clients can obtain IP addresses.

Configuration procedure

```
# Enable DHCP snooping.
<SwitchB> system-view
[SwitchB] dhcp-snooping

# Specify GigabitEthernet 1/0/1 as trusted.
[SwitchB] interface GigabitEthernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] dhcp-snooping trust
[SwitchB-GigabitEthernet1/0/1] quit

# Configure GigabitEthernet 1/0/2 to support Option 82.
[SwitchB] interface GigabitEthernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] dhcp-snooping information enable
[SwitchB-GigabitEthernet1/0/2] dhcp-snooping information strategy replace
```

```
[SwitchB-GigabitEthernet1/0/2] dhcp-snooping information circuit-id string company001
[SwitchB-GigabitEthernet1/0/2] dhcp-snooping information remote-id string device001
[SwitchB-GigabitEthernet1/0/2] quit
```

Configure GigabitEthernet 1/0/3 to support Option 82.

```
[SwitchB] interface GigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] dhcp-snooping information enable
[SwitchB-GigabitEthernet1/0/3] dhcp-snooping information strategy replace
[SwitchB-GigabitEthernet1/0/3] dhcp-snooping information format verbose node-identifier
sysname
[SwitchB-GigabitEthernet1/0/3] dhcp-snooping information circuit-id format-type ascii
[SwitchB-GigabitEthernet1/0/3] dhcp-snooping information remote-id format-type ascii
```

Configuring BOOTP client

Overview

BOOTP application

After you specify an interface of a device as a BOOTP client, the interface can use BOOTP to get information (such as IP address) from the BOOTP server.

To use BOOTP, an administrator must configure a BOOTP parameter file for each BOOTP client on the BOOTP server. The parameter file contains information such as MAC address and IP address of a BOOTP client. When a BOOTP client sends a request to the BOOTP server, the BOOTP server searches for the BOOTP parameter file and returns the corresponding configuration information.

BOOTP is usually used in relatively stable environments. In network environments that change frequently, DHCP is more suitable.

Because a DHCP server can interact with a BOOTP client, you can use the DHCP server to configure an IP address for the BOOTP client, without any BOOTP server.

Obtaining an IP address dynamically

A BOOTP client dynamically obtains an IP address from a BOOTP server in the following steps:

1. The BOOTP client broadcasts a BOOTP request, which contains its own MAC address.
2. The BOOTP server receives the request and searches the configuration file for the corresponding IP address and other information according to the MAC address of the BOOTP client. The BOOTP server then returns a BOOTP response to the BOOTP client.
3. The BOOTP client obtains the IP address from the received response.

A DHCP server can take the place of the BOOTP server in the above mentioned dynamic IP address acquisition.

Protocols and standards

- RFC 951, *Bootstrap Protocol (BOOTP)*
- RFC 2132, *DHCP Options and BOOTP Vendor Extensions*
- RFC 1542, *Clarifications and Extensions for the Bootstrap Protocol*

Configuration restrictions

- BOOTP client configuration only applies to Layer 3 Ethernet ports, Layer 3 aggregate interfaces and VLAN interfaces.
- If several VLAN interfaces sharing the same MAC address obtain IP addresses through a BOOTP relay agent, the BOOTP server cannot be a Windows Server 2000 or Windows Server 2003.
- You cannot configure an interface of an aggregation group as a BOOTP client.

Configuring an interface to dynamically obtain an IP address through BOOTP

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure an interface to dynamically obtain an IP address through BOOTP.	ip address bootp-alloc	By default, an interface does not use BOOTP to obtain an IP address.

Displaying and maintaining BOOTP client configuration

Task	Command	Remarks
Display BOOTP client information.	display bootp client [interface <i>interface-type</i> <i>interface-number</i>] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view

BOOTP client configuration example

Network requirements

As shown in [Figure 30](#), Switch B's port belonging to VLAN 1 is connected to the LAN. VLAN-interface 1 obtains an IP address from the DHCP server by using BOOTP.

Configuration procedure

The following describes only the configuration on Switch B serving as a client.

Configure VLAN-interface 1 to dynamically obtain an IP address from the DHCP server.

```
<SwitchB> system-view
[SwitchB] interface vlan-interface 1
[SwitchB-Vlan-interface1] ip address bootp-alloc
```

Use the **display bootp client** command to view the IP address assigned to the BOOTP client.

To make the BOOTP client obtain an IP address from the DHCP server, you must perform additional configurations on the DHCP server. For more information, see "[Configuring DHCP server.](#)"

Configuring IPv4 DNS

Overview

Domain Name System (DNS) is a distributed database used by TCP/IP applications to translate domain names into corresponding IP addresses. With DNS, you can use easy-to-remember domain names in some applications and let the DNS server translate them into correct IP addresses.

DNS services can be static or dynamic. After a user specifies a name, the device checks the local static name resolution table for an IP address. If no IP address is available, it contacts the DNS server for dynamic name resolution, which takes more time than static name resolution. To improve efficiency, you can put frequently queried name-to-IP address mappings in the local static name resolution table.

Static domain name resolution

Static domain name resolution means setting up mappings between domain names and IP addresses. IP addresses of the corresponding domain names can be found in the static domain resolution table when you use applications such as Telnet.

Dynamic domain name resolution

1. A user program sends a name query to the resolver of the DNS client.
2. The DNS resolver looks up the local domain name cache for a match. If the resolver finds a match, it sends the corresponding IP address back. If not, it sends a query to the DNS server.
3. The DNS server looks up the corresponding IP address of the domain name in its DNS database. If no match is found, the server sends a query to a higher level DNS server. This process continues until a result, whether successful or not, is returned.
4. After receiving a response from the DNS server, the DNS client returns the resolution result to the application.

Figure 39 Dynamic domain name resolution

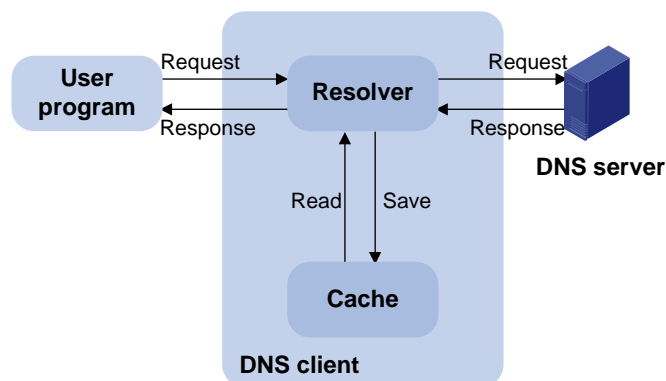


Figure 39 shows the relationship between the user program, DNS client, and DNS server.

The DNS client is made up of the resolver and cache. The user program and DNS client can run on the same device or different devices, but the DNS server and the DNS client usually run on different devices.

Dynamic domain name resolution allows the DNS client to store latest mappings between domain names and IP addresses in the dynamic domain name cache. The DNS client does not need to send a request to the DNS server for a repeated query next time. The aged mappings are removed from the cache after some time, and latest entries are required from the DNS server. The DNS server decides how long a mapping is valid, and the DNS client gets the aging information from DNS messages.

DNS suffixes

The DNS client holds a list of suffixes which the user sets. The resolver can use the list to supply the missing part of incomplete names.

For example, a user can configure com as the suffix for aabbcc.com. The user only needs to type aabbcc to obtain the IP address of aabbcc.com because the resolver adds the suffix and delimiter before passing the name to the DNS server.

- If there is no dot (.) in the domain name (for example, aabbcc), the resolver considers this a host name and adds a DNS suffix before the query. If no match is found after all the configured suffixes are used, the original domain name (for example, aabbcc) is used for the query.
- If there is a dot (.) in the domain name (for example, www.aabbcc), the resolver directly uses this domain name for the query. If the query fails, the resolver adds a DNS suffix for another query.
- If the dot (.) is at the end of the domain name (for example, aabbcc.com.), the resolver considers it a Fully Qualified Domain Name (FQDN) and returns the query result, successful or failed. The dot (.) is considered a terminating symbol.

The device supports static and dynamic DNS client services.

NOTE:

If an alias is configured for a domain name on the DNS server, the device can resolve the alias into the IP address of the host.

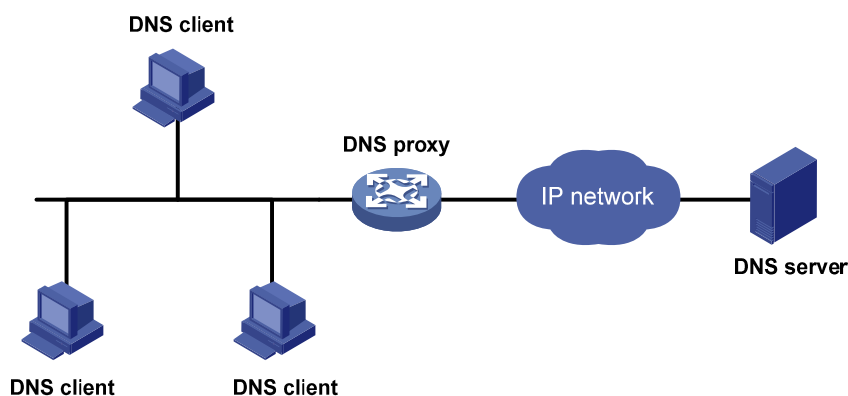
DNS proxy

A DNS proxy forwards DNS requests and replies between DNS clients and a DNS server.

As shown in [Figure 40](#), a DNS client sends a DNS request to the DNS proxy, which forwards the request to the designated DNS server, and conveys the reply from the DNS server to the client.

The DNS proxy simplifies network management. When the DNS server address is changed, you can change the configuration on only the DNS proxy instead of on each DNS client.

Figure 40 DNS proxy networking application



A DNS proxy operates as follows:

1. A DNS client considers the DNS proxy as the DNS server, and sends a DNS request to the DNS proxy. The destination address of the request is the IP address of the DNS proxy.
2. The DNS proxy searches the local static domain name resolution table and dynamic domain name resolution table after receiving the request. If the requested information is found, the DNS proxy returns a DNS reply to the client.
3. If the requested information is not found, the DNS proxy sends the request to the designated DNS server for domain name resolution.
4. After receiving a reply from the DNS server, the DNS proxy records the IP address-to-domain name mapping and forwards the reply to the DNS client.

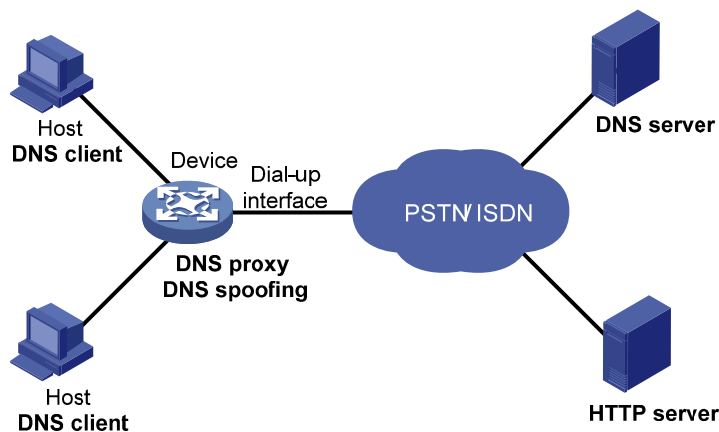
With no DNS server or route to a DNS server specified, the DNS proxy does not forward DNS requests, or answer requests from the DNS clients.

DNS spoofing

DNS spoofing is applied to the dial-up network, as shown in [Figure 41](#).

- The device connects to the PSTN/ISDN network through a dial-up interface and triggers the establishment of a dial-up connection only when packets are to be forwarded through the dial-up interface.
- The device serves as a DNS proxy and is specified as a DNS server on the hosts. After the dial-up connection is established through the dial-up interface, the device dynamically obtains the DNS server address through DHCP or other autoconfiguration mechanisms.

Figure 41 Application of DNS spoofing



Without DNS spoofing enabled, the device forwards the DNS requests received from the hosts to the DNS server, if it cannot find a match in the local domain name resolution table. However, without any dial-up connection established, the device cannot obtain the DNS server address, so it cannot forward or answer the requests from the clients. The domain name cannot be resolved and no traffic triggers the establishment of a dial-up connection.

DNS spoofing can solve this problem. DNS spoofing enables the device to reply the DNS client with a configured IP address when the device does not have a DNS server address or route to a DNS server. Subsequent packets sent by the DNS client trigger the establishment of a dial-up connection with the network.

In the network of [Figure 41](#), a host accesses the HTTP server in following these steps:

1. The host sends a DNS request to the device to resolve the domain name of the HTTP server into an IP address.
2. Upon receiving the request, the device searches the local static and dynamic DNS entries for a match. If no match is found and the device does know the DNS server address, the device spoofs the host by replying a configured IP address. The TTL of the DNS reply is 0. The device must have a route to the IP address with the dial-up interface as the outgoing interface.
3. Upon receiving the reply, the host sends an HTTP request to the replied IP address.
4. When forwarding the HTTP request through the dial-up interface, the device establishes a dial-up connection with the network and dynamically obtains the DNS server address through DHCP or other autoconfiguration mechanisms.
5. When the DNS reply ages out, the host sends a DNS request to the device again.
6. Then the device operates the same as a DNS proxy. For more information, see "[A DNS proxy operates as follows:](#)"
7. After obtaining the IP address of the HTTP server, the host can access the HTTP server.

Because the IP address configured with DNS spoofing is not the actual IP address of the requested domain name, the TTL of the DNS reply is set to 0 to prevent the DNS client from generating incorrect domain name-to-IP address mappings.

Configuring the IPv4 DNS client

Configuring static domain name resolution

Configuring static domain name resolution refers to specifying the mappings between host names and IPv4 addresses. Static domain name resolution allows applications such as Telnet to contact hosts by using host names instead of IPv4 addresses.

Follow these guidelines when you configure static domain name resolution:

- The IPv4 address you last assign to the host name will overwrite the previous one if there is any.
- You may create up to 50 static mappings between domain names and IPv4 addresses.

To configure static domain name resolution:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure a mapping between a host name and an IPv4 address.	ip host <i>hostname ip-address</i>	Not configured by default

Configuring dynamic domain name resolution

To send DNS queries to a correct server for resolution, dynamic domain name resolution needs to be enabled and a DNS server needs to be configured.

In addition, you can configure a DNS suffix that the system will automatically add to the provided domain name for resolution.

Configuration restrictions and guidelines

- You can configure up to six DNS servers, including those with IPv6 addresses, in system view, and up to six DNS servers on all interfaces of a device.
- A DNS server configured in system view has a higher priority than one configured in interface view. A DNS server configured earlier has a higher priority than one configured later in the same view. A DNS server manually configured has a higher priority than one dynamically obtained through DHCP. A name query request is first sent to the DNS server that has the highest priority. If no reply is received, it is sent to the DNS server that has the second highest priority, and thus in turn.
- You can specify up to ten DNS suffixes.

Configuration procedure

To configure dynamic domain name resolution:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable dynamic domain name resolution.	dns resolve	Disabled by default.
3. Specify a DNS server.	<ul style="list-style-type: none">• (Approach 1) In System view: dns server ip-address• (Approach 2) In interface view:<ul style="list-style-type: none">a. interface interface-type interface-numberb. dns server ip-addressc. quit	Use either approach. Not specified by default.
4. Configure a DNS suffix.	dns domain domain-name	Optional. Not configured by default. Only the provided domain name is resolved.

Configuring the DNS proxy

You can specify multiple DNS servers by using the **dns server** command repeatedly. Upon receiving a name query request from a client, the DNS proxy forwards the request to the DNS server that has the highest priority. If having not received a reply, it forwards the request to a DNS server that has the second highest priority, and thus in turn.

To configure the DNS proxy:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable DNS proxy.	dns proxy enable	Disabled by default.

Step	Command	Remarks
3. Specify a DNS server.	<ul style="list-style-type: none"> (Approach 1) In system view: dns server <i>ip-address</i> (Approach 2) In interface view: <ul style="list-style-type: none"> a. interface <i>interface-type</i> <i>interface-number</i> b. dns server <i>ip-address</i> 	<p>Use either approach.</p> <p>No DNS server is specified by default.</p>

Configuring DNS spoofing

DNS spoofing is effective only when:

- The DNS proxy is enabled on the device.
- No DNS server or route to any DNS server is specified on the device.

To configure DNS spoofing:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable DNS spoofing and specify the translated IP address.	dns spoofing <i>ip-address</i>	Disabled by default

Setting the DSCP value for DNS packets

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the DSCP value for DNS packets.	dns dscp <i>dscp-value</i>	<p>Optional.</p> <p>By default, the DSCP value for DNS packets is 0.</p>

Specifying the source interface for DNS packets

By default, the device uses the primary IP address of the output interface of the matching route as the source IP address of a DNS request. Therefore, the source IP address of the DNS packets may vary with DNS servers. In some scenarios, the DNS server only responds to DNS requests sourced from a specific IP address. In such cases, you must specify the source interface for the DNS packets so that the device can always use the primary IP address of the specified source interface as the source IP address of DNS packets.

To specify the source interface for DNS packets:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Set the DSCP value for DNS packets.	dns source-interface <i>interface-type interface-number</i>	By default, no source interface for DNS packets is specified. The device uses the primary IP address of the output interface of the matching route as the source IP address of a DNS request.

Displaying and maintaining IPv4 DNS

Task	Command	Remarks
Display the static IPv4 domain name resolution table.	display ip host [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv4 DNS server information.	display dns server [dynamic] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display DNS suffixes.	display dns domain [dynamic] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the information of the dynamic IPv4 domain name cache.	display dns host ip [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear the information of the dynamic IPv4 domain name cache.	reset dns host ip	Available in user view

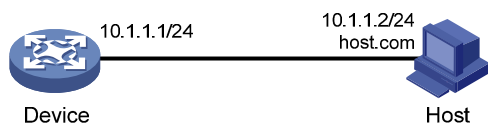
Static domain name resolution configuration example

Network requirements

As shown in [Figure 42](#), the device wants to access the host by using an easy-to-remember domain name rather than an IP address.

Configure static domain name resolution on the device so that the device can use the domain name `host.com` to access the host whose IP address is `10.1.1.2`.

Figure 42 Network diagram



Configuration procedure

Configure a mapping between host name `host.com` and IP address `10.1.1.2`.

```

<Sysname> system-view
[Sysname] ip host host.com 10.1.1.2
# Use the ping host.com command to verify that the device can use static domain name resolution to
resolve domain name host.com into IP address 10.1.1.2.
[Sysname] ping host.com
  PING host.com (10.1.1.2):
  56 data bytes, press CTRL_C to break
    Reply from 10.1.1.2: bytes=56 Sequence=1 ttl=128 time=1 ms
    Reply from 10.1.1.2: bytes=56 Sequence=2 ttl=128 time=4 ms
    Reply from 10.1.1.2: bytes=56 Sequence=3 ttl=128 time=3 ms
    Reply from 10.1.1.2: bytes=56 Sequence=4 ttl=128 time=2 ms
    Reply from 10.1.1.2: bytes=56 Sequence=5 ttl=128 time=3 ms

--- host.com ping statistics ---
  5 packet(s) transmitted
  5 packet(s) received
  0.00% packet loss
  round-trip min/avg/max = 1/2/4 ms

```

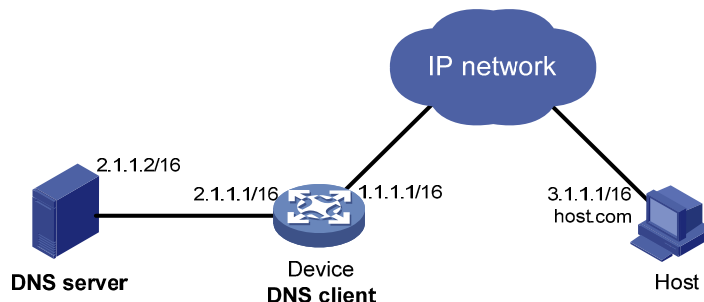
Dynamic domain name resolution configuration example

Network requirements

As shown in [Figure 43](#), the device wants to access the host by using an easy-to-remember domain name rather than an IP address, and to request the DNS server on the network for an IP address by using dynamic domain name resolution. The IP address of the DNS server is 2.1.1.2/16 and the DNS server has a com domain, which stores the mapping between domain name host and IP address 3.1.1.1/16.

Configure dynamic domain name resolution and the domain name suffix com on the device that serves as a DNS client so that the device can use domain name host to access the host with the domain name host.com and the IP address 3.1.1.1/16.

Figure 43 Network diagram



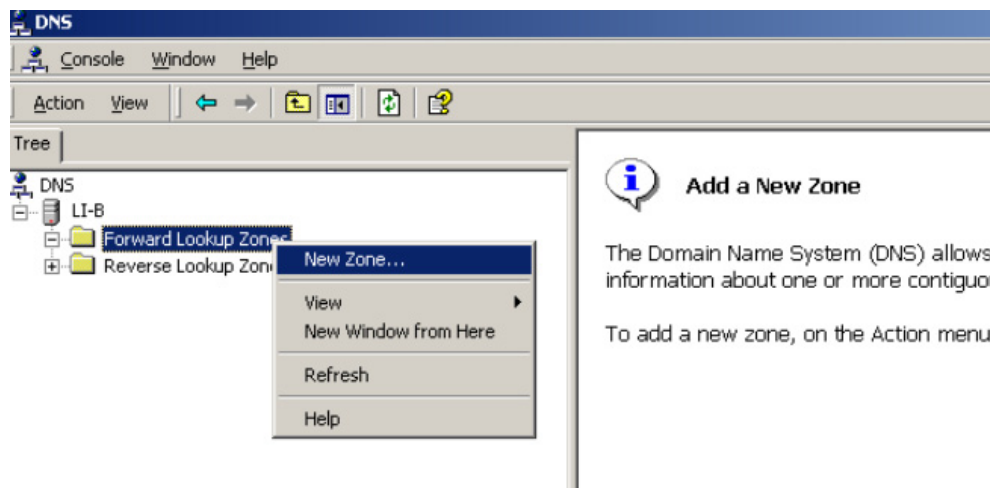
Configuration procedure

Before performing the following configuration, make sure that the device and the host are accessible to each other via available routes, and that the IP addresses of the interfaces are configured as shown [Figure 43](#).

This configuration may vary with DNS servers. The following configuration is performed on a PC running Windows Server 2000.

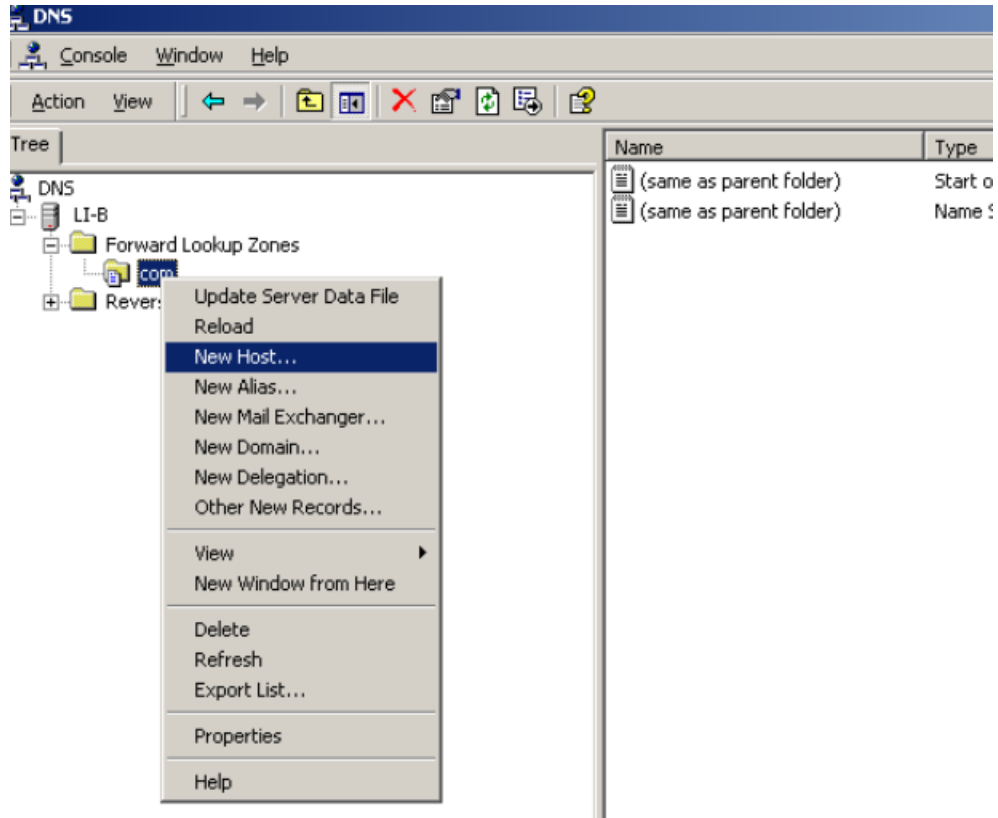
1. Configure the DNS server:
 - a. Select **Start > Programs > Administrative Tools > DNS**.
The DNS server configuration page appears, as shown in [Figure 44](#).
 - b. Right click **Forward Lookup Zones**, select **New Zone**, and then follow the steps to create a new zone named **com**.

Figure 44 Creating a zone



- a. On the DNS server configuration page, right click zone **com**, and select **New Host**.

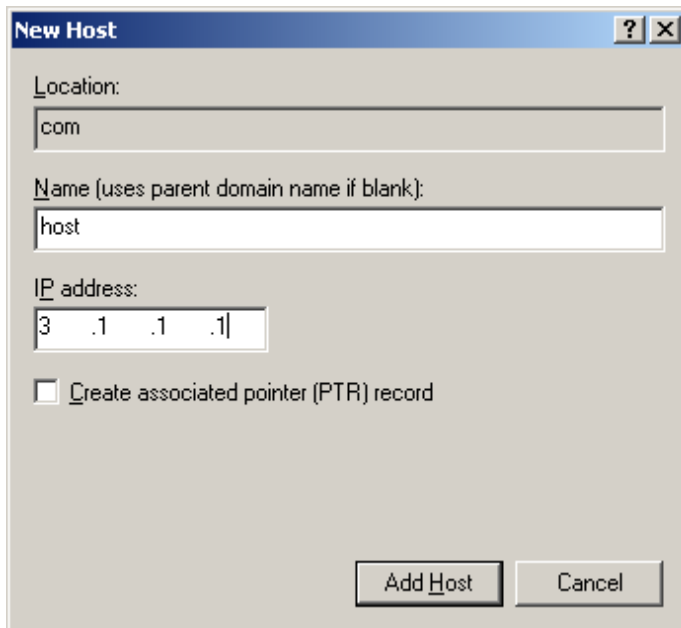
Figure 45 Adding a host



- d. On the page that appears, enter host name **host** and IP address **3.1.1.1**.
- e. Click **Add Host**.

The mapping between the IP address and host name is created.

Figure 46 Adding a mapping between domain name and IP address



- 2. Configure the DNS client:


```
# Enable dynamic domain name resolution.
<Sysname> system-view
[Sysname] dns resolve
# Specify the DNS server 2.1.1.2.
[Sysname] dns server 2.1.1.2
# Configure com as the name suffix.
[Sysname] dns domain com
```

Verifying the configuration

Use the **ping host** command on the device to verify that the communication between the device and the host is normal and that the corresponding destination IP address is 3.1.1.1.

```
[Sysname] ping host
Trying DNS resolve, press CTRL_C to break
Trying DNS server (2.1.1.2)
PING host.com (3.1.1.1):
56 data bytes, press CTRL_C to break
  Reply from 3.1.1.1: bytes=56 Sequence=1 ttl=126 time=3 ms
  Reply from 3.1.1.1: bytes=56 Sequence=2 ttl=126 time=1 ms
  Reply from 3.1.1.1: bytes=56 Sequence=3 ttl=126 time=1 ms
  Reply from 3.1.1.1: bytes=56 Sequence=4 ttl=126 time=1 ms
  Reply from 3.1.1.1: bytes=56 Sequence=5 ttl=126 time=1 ms

--- host.com ping statistics ---
 5 packet(s) transmitted
 5 packet(s) received
 0.00% packet loss
 round-trip min/avg/max = 1/1/3 ms
```

DNS proxy configuration example

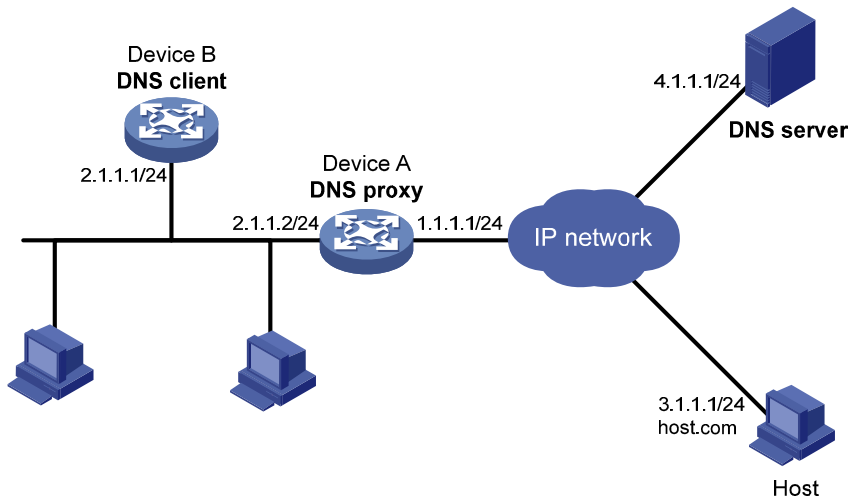
Network requirements

When the IP address of the DNS server changes, you must configure the new IP address of the DNS server on each device on the LAN. To simplify network management, you can use the DNS proxy function.

As shown in [Figure 47](#):

- Specify Device A as the DNS server of Device B (the DNS client). Device A acts as a DNS proxy. The IP address of the real DNS server is 4.1.1.1.
- Configure the IP address of the DNS proxy on Device B. DNS requests of Device B are forwarded to the real DNS server through the DNS proxy.

Figure 47 Network diagram



Configuration procedure

Before performing the following configuration, assume that Device A, the DNS server, and the host are reachable to each other and the IP addresses of the interfaces are configured as shown in [Figure 47](#).

1. Configure the DNS server:
This configuration may vary with different DNS servers. When a PC running Windows Server 2000 acts as the DNS server, see "[Dynamic domain name resolution configuration example](#)" for related configuration information.

2. Configure the DNS proxy:
Specify the DNS server 4.1.1.1.
<DeviceA> system-view
[DeviceA] dns server 4.1.1.1
Enable DNS proxy.
[DeviceA] dns proxy enable

3. Configure the DNS client:
Enable the domain name resolution function.
<DeviceB> system-view
[DeviceB] dns resolve
Specify the DNS server 2.1.1.2.
[DeviceB] dns server 2.1.1.2

Verifying the configuration

Execute the **ping host.com** command on Device B to verify that the communication between the device and the host is normal and that the corresponding destination IP address is 3.1.1.1.

```
[DeviceB] ping host.com
Trying DNS resolve, press CTRL_C to break
Trying DNS server (2.1.1.2)
PING host.com (3.1.1.1):
56 data bytes, press CTRL_C to break
```

```
Reply from 3.1.1.1: bytes=56 Sequence=1 ttl=126 time=3 ms
Reply from 3.1.1.1: bytes=56 Sequence=2 ttl=126 time=1 ms
Reply from 3.1.1.1: bytes=56 Sequence=3 ttl=126 time=1 ms
Reply from 3.1.1.1: bytes=56 Sequence=4 ttl=126 time=1 ms
Reply from 3.1.1.1: bytes=56 Sequence=5 ttl=126 time=1 ms

--- host.com ping statistics ---
 5 packet(s) transmitted
 5 packet(s) received
 0.00% packet loss
round-trip min/avg/max = 1/1/3 ms
```

Troubleshooting IPv4 DNS configuration

Symptom

After enabling dynamic domain name resolution, the user cannot get the correct IP address.

Solution

1. Use the **display dns host ip** command to verify that the specified domain name is in the cache.
2. If the specified domain name does not exist, verify that dynamic domain name resolution is enabled and that the DNS client can communicate with the DNS server.
3. If the specified domain name is in the cache, but the IP address is incorrect, verify that the DNS client has the correct IP address of the DNS server.
4. Verify that the mapping between the domain name and IP address is correct on the DNS server.

Configuring IRDP

Overview

As an extension of the Internet Control Message Protocol (ICMP), the ICMP Router Discovery Protocol (IRDP) enables hosts to discover the IP addresses of their neighboring routers and set their default routes.

NOTE:

The hosts in this chapter support IRDP.

Background

Before a host can send packets to another network, it must know the IP address of at least one router on the local subnet. The host can obtain this information either through manual configuration, or from routing protocol packets sent by routers on the local subnet.

Both methods have disadvantages. The first method requires the administrator to manually configure and maintain router address information on hosts, and cannot track dynamic changes. The second method requires hosts to recognize various routing protocols, and will fail to work if no routing protocol runs on the local subnet.

IRDP was introduced to solve the problem. IRDP uses two new types of ICMP messages to allow hosts to discover neighboring routers. IRDP adapts to dynamic changes, requires less manual configuration, and does not rely on any routing protocols.

Working mechanism

IRDP uses the following types of ICMP messages.

- **Router advertisement (RA)**—Sent by a router to advertise its IP address and preference.
- **Router solicitation (RS)**—Sent by a host to voluntarily request the IP addresses of routers on the subnet.

IRDP operates in the following steps:

1. A router periodically broadcasts or multicasts an RA, which contains the IP addresses (including the primary IP address and manually configured secondary IP addresses) of interfaces. Hosts listen for RAs to obtain the IP addresses of neighboring routers.
2. Rather than wait for RAs, a newly attached host can voluntarily send an RS to request immediate RAs for the IP addresses of routers on the subnet. If no response to the RS is received, the host retransmits the RS several times. If the host still receives no RAs, it will obtain the IP addresses of routers from periodic RAs.
3. Upon receiving an RA, a host adds the IP addresses in the RA to its routing table. The host selects the IP address with the highest preference among all obtained IP addresses as the default gateway.

IRDP allows hosts to locate routers, but does not suggest the best route to a specific destination. If a host selects a router that is not the best next hop to a specific destination, the router will send back an ICMP redirect message to provide a better next hop.

Concepts

Preference of an IP address

Every IP address advertised in RAs has a preference value. The IP address with the highest preference is selected as the default router address.

You can configure the preference for IP addresses advertised on a router interface.

The bigger the preference value, the higher the preference. The minimum preference value (-2147483648) is used to indicate that the address, even though it may be advertised, is not to be used by neighboring hosts as a default router address.

Lifetime of an IP address

An RA contains a lifetime field that specifies the lifetime of advertised IP addresses. If no new RA for an IP address is received within the lifetime of the IP address, the host removes the corresponding route information.

All the IP addresses advertised by an interface have the same lifetime.

Advertising interval

A router interface with IRDP enabled sends out RAs at a random interval between the minimum advertising interval and the maximum advertising interval. This mechanism prevents the local link from being overloaded by a large number of RAs sent simultaneously from routers.

HP recommends shortening the advertising interval on a link that suffers high packet loss rates.

Destination address of RAs

An RA uses either of the two destination IP addresses:

- broadcast address 255.255.255.255.
- Multicast address 224.0.0.1, which identifies all the hosts on the local subnet.

By default, the destination IP address of an RA is the broadcast address. If the interface that sends RAs supports multicast, configure 224.0.0.1 as the destination IP address.

Proxy-advertised IP addresses

By default, an interface advertises its primary IP address and manually configured secondary IP addresses. You can configure other IP addresses for an interface to proxy-advertise.

Protocols and standards

RFC 1256, *ICMP Router Discovery Messages*

Configuration procedure

IRDP configuration takes effect only when IRDP is enabled.

To configure IRDP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	The interface can be a Layer 3 Ethernet port or VLAN interface.
3. Enable IRDP on the interface.	ip irdp	Disabled by default.
4. Configure the preference of advertised IP addresses.	ip irdp preference <i>preference-value</i>	Optional. The preference defaults to 0. The specified preference applies to all advertised IP addresses, including the primary IP address and the manually configured secondary IP addresses of the interface.
5. Set the lifetime of advertised IP addresses.	ip irdp lifetime <i>life-number</i>	Optional. 1800 seconds by default. The specified lifetime applies to all advertised IP addresses, including the IP address of the interface and proxy-advertised IP addresses on the interface.
6. Set the minimum advertising interval.	ip irdp minadvinterval <i>min-value</i>	Optional. 450 seconds by default.
7. Set the maximum advertising interval.	ip irdp maxadvinterval <i>max-value</i>	Optional. 600 seconds by default.
8. Configure the multicast address (224.0.0.1) as the destination IP address of RAs.	ip irdp multicast	Optional. By default, RAs use the broadcast address 255.255.255.255 as the destination IP address.
9. Specify a proxy-advertised IP address and its preference.	ip irdp address <i>ip-address</i> <i>preference</i>	Optional.

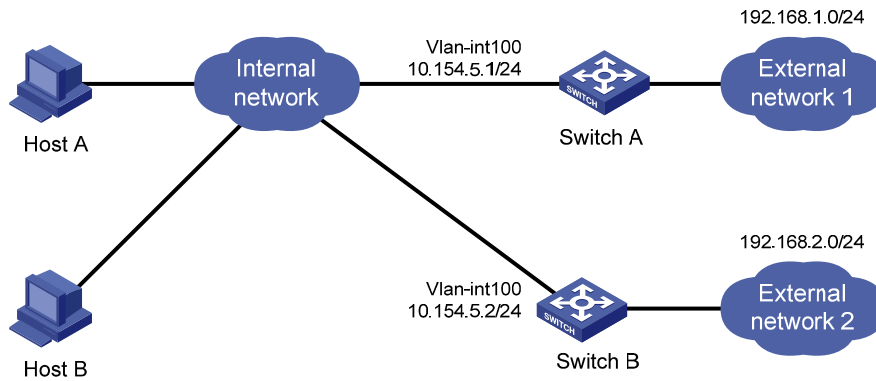
IRDP configuration example

Network requirements

Host A and Host B that run Linux operating systems reside in the internal network of a company. Switch A and Switch B serve as the egress routers and connect to external networks 192.168.1.0/24 and 192.168.2.0/24 respectively.

Configure Switch A as the default gateway of the hosts. The packets to the external networks can be properly routed.

Figure 48 Network diagram



Configuration procedure

1. Configure Switch A:

Specify the IP address for VLAN-interface 100.

```
<SwitchA> system-view
[SwitchA] interface Vlan-interface 100
[SwitchA-Vlan-interface100] ip address 10.154.5.1 24
```

Enable IRDP on VLAN-interface 100.

```
[SwitchA-Vlan-interface100] ip irdp
```

Specify preference 1000 for the IP address of VLAN-interface 100.

```
[SwitchA-Vlan-interface100] ip irdp preference 1000
```

Configure the multicast address 224.0.0.1 as the destination IP address for RAs sent by VLAN-interface 100.

```
[SwitchA-Vlan-interface100] ip irdp multicast
```

Specify the IP address 192.168.1.0 and preference 400 for VLAN-interface 100 to proxy-advertise.

```
[SwitchA-Vlan-interface100] ip irdp address 192.168.1.0 400
```

2. Configure Switch B:

Specify the IP address of VLAN-interface 100.

```
<SwitchB> system-view
[SwitchB] interface Vlan-interface 100
[SwitchB-Vlan-interface100] ip address 10.154.5.2 24
```

Enable IRDP on VLAN-interface 100.

```
[SwitchB-Vlan-interface100] ip irdp
```

Specify preference 500 for the IP address of VLAN-interface 100.

```
[SwitchB-Vlan-interface100] ip irdp preference 500
```

Configure the multicast address 224.0.0.1 as the destination IP address for RAs sent by VLAN-interface 100.

```
[SwitchB-Vlan-interface100] ip irdp multicast
```

Specify the IP address 192.168.2.0 and preference 400 for VLAN-interface 100 to proxy-advertise.

```
[SwitchB-Vlan-interface100] ip irdp address 192.168.2.0 400
```

Verifying the configuration

After enabling IRDP on Host A and Host B, display the routing table for the hosts (Host A for example).

```
[HostA@localhost ~]$ netstat -rne
```

```
Kernel IP routing table
```

Destination	Gateway	Genmask	Flags	Metric	Ref	Use	Iface
10.154.5.0	0.0.0.0	255.255.255.0	U	0	0	0	eth1
192.168.1.0	0.0.0.0	255.255.255.0	U	0	0	0	eth1
192.168.2.0	0.0.0.0	255.255.255.0	U	0	0	0	eth1
0.0.0.0	10.154.5.1	0.0.0.0	UG	0	0	0	eth1

The output shows that the default route on Host A points to IP address 10.154.5.1, and Host A has routes to 192.168.1.0/24 and 192.168.2.0/24.

Configuring IP performance optimization

Only the HP 5500 EI switch series supports configuring IP performance optimization on Layer 3 Ethernet interfaces.

The term "interface" in this chapter refers to Layer 3 interfaces, including VLAN interfaces and route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).

Enabling receiving and forwarding of directed broadcasts to a directly connected network

Directed broadcast packets are broadcast on a specific network. In the destination IP address of a directed broadcast, the network ID identifies the target network, and the host ID is made up of all ones. If a device is allowed to forward directed broadcasts to a directly connected network, hackers may mount attacks to the network. However, you can enable the feature by using the UDP Helper function to convert broadcasts to unicasts and forward them to a specified server.

Enabling receiving of directed broadcasts to a directly connected network

If the switch is enabled to receive directed broadcasts, the switch determines whether to forward them according to the configuration on the outgoing interface.

To enable the device to receive directed broadcasts:

Step	Command
1. Enter system view.	system-view
2. Enable the device to receive directed broadcasts.	ip forward-broadcast

Enabling forwarding of directed broadcasts to a directly connected network

Follow these guidelines when you enable the device to forward directed broadcasts:

- If an ACL is referenced in the **ip forward-broadcast** command, only packets permitted by the ACL can be forwarded.
- If you repeatedly execute the **ip forward-broadcast** command on an interface, only the last command takes effect. If the command executed last does not include **acl acl-number**, the ACL configured previously is removed.

To enable the device to forward directed broadcasts:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type interface-number</i>	N/A
3. Enable the interface to forward directed broadcasts.	ip forward-broadcast [acl <i>acl-number</i>]	Disabled by default

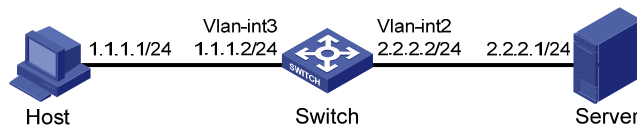
Configuration example

Network requirements

As shown in Figure 49, the host's interface and VLAN-interface 3 of the switch are on the same network segment (1.1.1.0/24). VLAN-interface 2 of Switch and the server are on another network segment (2.2.2.0/24). The default gateway of the host is VLAN-interface 3 (IP address 1.1.1.2/24) of Switch.

Configure the switch so that the server can receive directed broadcasts from the host to IP address 2.2.2.255.

Figure 49 Network diagram



Configuration procedure

Enable the switch to receive directed broadcasts.

```
<Switch> system-view
[Switch] ip forward-broadcast
```

Configure IP addresses for VLAN-interface 3 and VLAN-interface 2.

```
[Switch] interface vlan-interface 3
[Switch-Vlan-interface3] ip address 1.1.1.2 24
[Switch-Vlan-interface3] quit
[Switch] interface vlan-interface 2
[Switch-Vlan-interface2] ip address 2.2.2.2 24
```

Enable VLAN-interface 2 to forward directed broadcasts.

```
[Switch-Vlan-interface2] ip forward-broadcast
```

Configuring TCP attributes

Configuring TCP path MTU discovery

ⓘ IMPORTANT:

All the devices on the TCP path must be enabled to send ICMP error messages by using the **ip unreachable enable** command.

TCP path MTU discovery (in RFC 1191) discovers the path MTU between the source and destination ends of a TCP connection. It works as follows:

1. A TCP source device sends a packet with the Don't Fragment (DF) bit set.
2. A router that fails to forward the packet because it exceeds the MTU on the outgoing interface discards the packet and returns an ICMP error message, which contains the MTU of the outgoing interface.
3. Upon receiving the ICMP message, the TCP source device calculates the current path MTU of the TCP connection.
4. The TCP source device sends subsequent TCP segments that each are smaller than the MSS (MSS = path MTU - IP header length - TCP header length).

If the TCP source device still receives ICMP error messages when the MSS is smaller than 32 bytes, the TCP source device will fragment packets.

An ICMP error message received from a router that does not support RFC 1191 has the MTU of the outgoing interface set to 0. Upon receiving the ICMP message, the TCP source device selects the path MTU smaller than the current path MTU from the MTU table as described in RFC 1191 to calculate the TCP MSS. The MTU table contains MTUs of 68, 296, 508, 1006, 1280, 1492, 2002, 4352, 8166, 17914, 32000, and 65535 bytes. Because the minimum TCP MSS specified by the system is 32 bytes, the actual minimum MTU is 72 bytes.

After you enable TCP path MTU discovery, all new TCP connections will detect the path MTU. The device uses the path MTU to calculate the MSS to avoid IP fragmentation.

The path MTU uses an aging mechanism to make sure that the source device can increase the path MTU when the minimum link MTU on the path increases.

- When the TCP source device receives an ICMP error message, it reduces the path MTU and starts an age timer for the path MTU.
- After the age timer expires, the source device uses a larger MSS in the MTU table as described in RFC 1191.
- If no ICMP error message is received within two minutes, the source device increases the MSS again until the MSS is as large as the MSS negotiated during TCP three-way handshake.

To enable TCP path MTU discovery:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable TCP path MTU discovery.	tcp path-mtu-discovery [aging <i>minutes</i> no-aging]	Optional. Disabled by default.

Configuring the TCP send/receive buffer size

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the size of TCP send/receive buffer.	tcp window <i>window-size</i>	Optional. 8 KB by default.

Configuring TCP timers

You can configure the following TCP timers:

- **synwait timer**—When sending a SYN packet, TCP starts the synwait timer. If no response packet is received within the synwait timer interval, the TCP connection cannot be created.
- **finwait timer**—When a TCP connection is changed into FIN_WAIT_2 state, the finwait timer is started. If no FIN packet is received within the timer interval, the TCP connection is terminated. If a FIN packet is received, the TCP connection state changes to TIME_WAIT. If a non-FIN packet is received, the system restarts the timer upon receiving the last non-FIN packet. The connection is broken after the timer expires.

The actual length of the finwait timer is determined by the following formula:

Actual length of the finwait timer = (Configured length of the finwait timer – 75) + configured length of the synwait timer

To configure TCP timers:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the TCP synwait timer.	tcp timer syn-timeout <i>time-value</i>	Optional. 75 seconds by default.
3. Configure the TCP finwait timer.	tcp timer fin-timeout <i>time-value</i>	Optional. 675 seconds by default.

Configuring ICMP to send error packets

Sending error packets is a major function of ICMP. In case of network abnormalities, error packets are usually sent by the network or transport layer protocols to notify corresponding devices so as to facilitate control and management.

Advantages of sending ICMP error packets

ICMP error packets include the following types:

- ICMP redirect packets

A host may have only a default route to the default gateway in its routing table after startup. If the following conditions are satisfied, the default gateway will send ICMP redirect packets to the source host, telling it to reselect a correct next hop to send the subsequent packets:

- The receiving and forwarding interfaces are the same.
- The selected route has not been created or modified by an ICMP redirect packet.
- The selected route is not the default route of the device.
- There is no source route option in the packet.

The ICMP redirect packets function simplifies host administration and enables a host to gradually establish a sound routing table to find the best route.

- ICMP timeout packets

If the device receives an IP packet with a timeout error, it drops the packet and sends an ICMP timeout packet to the source.

The device sends an ICMP timeout packet under the following conditions:

- If the device finds that the destination of a packet is not itself and the TTL field of the packet is 1, it will send a "TTL timeout" ICMP error message.
- When the device receives the first fragment of an IP datagram whose destination is the device itself, it starts a timer. If the timer times out before all the fragments of the datagram are received, the device will send a "reassembly timeout" ICMP error packet.
- ICMP destination unreachable packets

If the device receives an IP packet with the destination unreachable, it will drop the packet and send an ICMP destination unreachable error packet to the source.

Conditions for sending an ICMP destination unreachable packet:

 - If neither a route nor the default route for forwarding a packet is available, the device will send a "network unreachable" ICMP error packet.
 - If the destination of a packet is local but the transport layer protocol of the packet is not supported by the local device, the device sends a "protocol unreachable" ICMP error packet to the source.
 - When receiving a packet with the destination being local and transport layer protocol being UDP, if the packet's port number does not match the running process, the device will send the source a "port unreachable" ICMP error packet.
 - If the source uses "strict source routing" to send packets, but the intermediate device finds that the next hop specified by the source is not directly connected, the device will send the source a "source routing failure" ICMP error packet.
 - When forwarding a packet, if the MTU of the sending interface is smaller than the packet, but the packet has been set as "Don't Fragment," the device will send the source a "fragmentation needed and Don't Fragment (DF)-set" ICMP error packet.

Disadvantages of sending ICMP error packets

Sending ICMP error packets facilitates network control and management, but it has the following disadvantages:

- Increases network traffic.
- A device's performance degrades if it receives a lot of malicious packets that cause it to respond with ICMP error packets.
- A host's performance degrades if the redirection function increases the size of its routing table.
- End users are affected because of receiving ICMP destination unreachable packets caused by malicious users.

To prevent such problems, disable the device from sending ICMP error packets.

Configuration procedure

The device stops sending "TTL timeout" ICMP error packets after sending ICMP timeout packets is disabled. However, "reassembly timeout" error packets will be sent normally.

To enable sending of ICMP error packets:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable sending of ICMP redirect packets.	ip redirects enable	Disabled by default

Step	Command	Remarks
3. Enable sending of ICMP timeout packets.	ip ttl-expires enable	Disabled by default
4. Enable sending of ICMP destination unreachable packets.	ip unreachable enable	Disabled by default

Displaying and maintaining IP performance optimization

Task	Command	Remarks
Display TCP connection statistics.	display tcp statistics [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display UDP statistics.	display udp statistics [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display statistics of IP packets.	display ip statistics [<i>slot slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display ICMP statistics.	display icmp statistics [<i>slot slot-number</i>] [{ begin exclude include } <i>regular-expression</i>] <i>regular-expression</i>]	Available in any view
Display socket information.	display ip socket [<i>socktype sock-type</i>] [<i>task-id socket-id</i>] [<i>slot slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display FIB information.	display fib [<i>vpn-instance vpn-instance-name</i>] [<i>acl acl-number</i> <i>ip-prefix ip-prefix-name</i>] [{ begin include exclude } <i>regular-expression</i>]	Available in any view Only HP 5500 EI Switch Series supports the vpn-instance <i>vpn-instance-name</i> option
Display FIB information matching the specified destination IP address.	display fib [<i>vpn-instance vpn-instance-name</i>] <i>ip-address</i> [<i>mask</i> <i>mask-length</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view Only HP 5500 EI Switch Series supports the vpn-instance <i>vpn-instance-name</i> option
Clear statistics of IP packets.	reset ip statistics [<i>slot slot-number</i>]	Available in user view
Clear statistics of TCP connections.	reset tcp statistics	Available in user view
Clear statistics of UDP traffic.	reset udp statistics	Available in user view

Configuring UDP helper

Only the HP 5500 EI switches support configuring UDP helper on Layer 3 Ethernet interfaces.

The term "interface" in this chapter refers to Layer 3 interfaces, including VLAN interfaces and route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).

Overview

UDP helper functions as a relay agent that converts UDP broadcast packets into unicast packets and forwards them to a specified destination server. This is helpful when a host cannot obtain network configuration information or request device names through broadcasting because the server or host to be requested is located on another broadcast domain.

With UDP helper enabled, a device decides whether to forward a received UDP broadcast packet according to the UDP destination port number of the packet.

- If the destination port number of the packet matches the one pre-configured on the device, the device modifies the destination IP address in the IP header, and then sends the packet to the specified destination server.
- If the destination port number of the packet does not match the one pre-configured on the device, the device sends the packet to the upper layer protocol for processing.

Configuration restrictions and guidelines

- The receiving of directed broadcasts to a directly connected network is disabled by default on the switch. As a result, UDP helper is available only when the **ip forward-broadcast** command is configured in system view. For more information about reception and forwarding of directed broadcasts to a directly connected network, see "Configuring IP performance optimization."
- A UDP helper enabled device must not forward DHCP broadcast packets that use destination port 67 or 68. Therefore, the UDP port numbers set with the **udp-helper port** command must not include 67 or 68.
- You can specify a port number or the corresponding parameter for a UDP port to forward packets. For example, **udp-helper port 53** and **udp-helper port dns** specify the same UDP port number.
- The configuration of all UDP ports is removed if you disable UDP helper.
- You can configure up to 256 UDP port numbers to enable the forwarding of packets with these UDP port numbers.
- You can configure up to 20 destination servers on an interface.

Configuration procedure

To configure UDP helper:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enable UDP helper.	udp-helper enable	Disabled by default.
3. Enable the forwarding of packets with the specified UDP destination port number(s).	udp-helper port { <i>port-number</i> dns netbios-ds netbios-ns tacacs ftpp time }	No UDP port number is specified by default.
4. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
5. Specify the destination server to which UDP packets are to be forwarded.	udp-helper server <i>ip-address</i>	No destination server is specified by default.

Displaying and maintaining UDP helper

Task	Command	Remarks
Displays information about forwarded UDP packets.	display udp-helper server [interface <i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear statistics about packets forwarded.	reset udp-helper packet	Available in user view

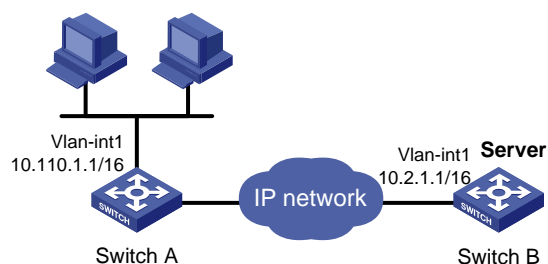
UDP helper configuration example

Network requirements

As shown in Figure 50, the IP address of VLAN-interface 1 of Switch A is 10.110.1.1/16, and the interface connects to the subnet 10.110.0.0/16.

Configure UDP helper to forward broadcast packets with UDP destination port number 55 and destination IP address 255.255.255.255 or 10.110.255.255 to the destination server 10.2.1.1/16 in public network.

Figure 50 Network diagram



Configuration procedure

Verify that a route from Switch A to the subnet 10.2.0.0/16 is available.


```
# Enable Switch A to receive directed broadcasts.
<SwitchA> system-view
[SwitchA] ip forward-broadcast

# Enable UDP helper.
[SwitchA] udp-helper enable

# Enable the forwarding broadcast packets with the UDP destination port 55.
[SwitchA] udp-helper port 55

# Specify the destination server 10.2.1.1 on VLAN-interface 1 in public network.
[SwitchA] interface vlan-interface 1
[SwitchA-Vlan-interfacel] ip address 10.110.1.1 16
[SwitchA-Vlan-interfacel] udp-helper server 10.2.1.1
```

Configuring IPv6 basics

Only the HP 5500 EI switches support configuring IPv6 basics on Layer 3 Ethernet interfaces.

The term "interface" in this chapter refers to Layer 3 interfaces, including VLAN interfaces and route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).

Overview

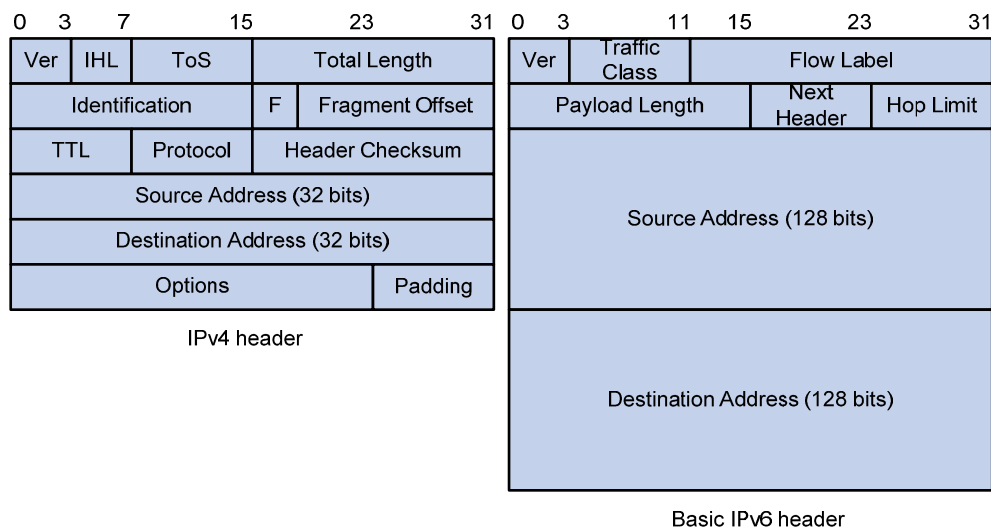
Internet Protocol Version 6 (IPv6), also called IP next generation (IPng), was designed by the Internet Engineering Task Force (IETF) as the successor to Internet Protocol version 4 (IPv4). The significant difference between IPv6 and IPv4 is that IPv6 increases the IP address size from 32 bits to 128 bits.

IPv6 features

Header format simplification

IPv6 removes several IPv4 header fields or moves them to the IPv6 extension headers to reduce the length of the basic IPv6 packet header. The basic IPv6 packet header has a fixed length of 40 bytes to simplify IPv6 packet handling and to improve forwarding efficiency. Although IPv6 address size is four times larger than IPv4 address size, the basic IPv6 packet header size is only twice the size of the option-less IPv4 packet header.

Figure 51 IPv4 packet header format and basic IPv6 packet header format



Larger address space

The source and destination IPv6 addresses are 128 bits (or 16 bytes) long. IPv6 can provide 3.4×10^{38} addresses to meet the requirements of hierarchical address division and the allocation of public and private addresses.

Hierarchical address structure

IPv6 uses hierarchical address structure to speed up route lookups and reduce the IPv6 routing table size through route aggregation.

Address autoconfiguration

To simplify host configuration, IPv6 supports stateful and stateless address autoconfiguration.

- Stateful address autoconfiguration enables a host to acquire an IPv6 address and other configuration information from a server (for example, a DHCP server).
- Stateless address autoconfiguration enables a host to automatically generate an IPv6 address and other configuration information by using its link-layer address and the prefix information advertised by a router.

To communicate with other hosts on the same link, a host automatically generates a link-local address based on its link-layer address and the link-local address prefix (FE80::/10).

Built-in security

IPv6 defines extension headers to support IPsec. IPsec provides end-to-end security for network security solutions and enhances interoperability among different IPv6 applications.

QoS support

The Flow Label field in the IPv6 header allows the device to label the packets and facilitates the special handling of a flow.

Enhanced neighbor discovery mechanism

The IPv6 neighbor discovery protocol is implemented through a group of Internet Control Message Protocol version 6 (ICMPv6) messages to manage the information exchange among neighboring nodes on the same link. The group of ICMPv6 messages replaces Address Resolution Protocol (ARP) messages, Internet Control Message Protocol version 4 (ICMPv4) Router Discovery messages, and ICMPv4 Redirect messages and provides a series of other functions.

Flexible extension headers

IPv6 eliminates the Options field in the header and introduces optional extension headers to provide scalability and improve efficiency. The Options field in the IPv4 packet header contains a maximum of 40 bytes, whereas the IPv6 extension headers are restricted to the maximum size of IPv6 packets only.

IPv6 addresses

IPv6 address format

An IPv6 address is represented as a set of 16-bit hexadecimal numbers separated by colons. An IPv6 address is divided into eight groups, and each 16-bit group is represented by four hexadecimal numbers, for example, 2001:0000:130F:0000:0000:09C0:876A:130B.

To simplify the representation of IPv6 addresses, you can handle zeros in IPv6 addresses by using the following methods:

- The leading zeros in each group can be removed. For example, the previous address can be represented in a shorter format as 2001:0:130F:0:0:9C0:876A:130B.
- If an IPv6 address contains two or more consecutive groups of zeros, they can be replaced by a double colon (::). For example, the previous address can be represented in the shortest format as 2001:0:130F::9C0:876A:130B.

A double colon may appear once or not at all in an IPv6 address. This limit allows the device to determine how many zeros the double colon represents, and correctly convert it to zeros to restore a 128-bit IPv6 address.

An IPv6 address consists of an address prefix and an interface ID, both of which are equivalent to the network ID and the host ID of an IPv4 address, respectively.

An IPv6 address prefix is written in IPv6-address/prefix-length notation where the IPv6-address is represented in any of the formats previously mentioned and the prefix-length is a decimal number indicating how many leftmost bits of the IPv6 address comprises the address prefix.

IPv6 address types

IPv6 addresses fall into the following types:

- **Unicast address**—An identifier for a single interface, similar to an IPv4 unicast address. A packet sent to a unicast address is delivered to the interface identified by that address.
- **Multicast address**—An identifier for a set of interfaces (typically belonging to different nodes), similar to an IPv4 multicast address. A packet sent to a multicast address is delivered to all interfaces identified by that address.
- **Anycast address**—An identifier for a set of interfaces (typically belonging to different nodes). A packet sent to an anycast address is delivered to the nearest one of the interfaces identified by that address. The nearest interface is chosen according to the routing protocols' measure of distance.

NOTE:

There are no broadcast addresses in IPv6. Their function is replaced by multicast addresses.

The type of an IPv6 address is designated by the first several bits, the format prefix. [Table 6](#) lists the mappings between address types and format prefixes.

Table 6 Mappings between address types and format prefixes

Type	Format prefix (binary)	IPv6 prefix ID	
Unicast address	Unspecified address	00...0 (128 bits)	::/128
	Loopback address	00...1 (128 bits)	::1/128
	Link-local address	1111111010	FE80::/10
	Site-local address	1111111011	FECO::/10
	Global unicast address	Other forms	N/A
Multicast address		11111111	FF00::/8
Anycast address	Anycast addresses use the unicast address space and have the identical structure of unicast addresses.		

Unicast addresses

Unicast addresses comprise global unicast addresses, link-local unicast addresses, site-local unicast addresses, the loopback address, and the unspecified address.

- Global unicast addresses, equivalent to public IPv4 addresses, are provided for network service providers. This type of address allows efficient prefix aggregation to restrict the number of global routing entries.

- Link-local addresses are used for communication among link-local nodes for neighbor discovery and stateless autoconfiguration. Packets with link-local source or destination addresses are not forwarded to other links.
- Site-local unicast addresses are similar to private IPv4 addresses. Packets with site-local source or destination addresses are not forwarded out of the local site (or a private network).
- A loopback address is 0:0:0:0:0:0:0:1 (or ::1). It cannot be assigned to any physical interface and can be used by a node to send an IPv6 packet to itself in the same way as the loopback address in IPv4.
- An unspecified address is 0:0:0:0:0:0:0:0 (or ::). It cannot be assigned to any node. Before acquiring a valid IPv6 address, a node fills this address in the source address field of IPv6 packets. The unspecified address cannot be used as a destination IPv6 address.

Multicast addresses

IPv6 multicast addresses listed in [Table 7](#) are reserved for special purposes.

Table 7 Reserved IPv6 multicast addresses

Address	Application
FF01::1	Node-local scope all-nodes multicast address
FF02::1	Link-local scope all-nodes multicast address
FF01::2	Node-local scope all-routers multicast address
FF02::2	Link-local scope all-routers multicast address
FF05::2	Site-local scope all-routers multicast address

Multicast addresses also include solicited-node addresses. A node uses a solicited-node multicast address to acquire the link-layer address of a neighboring node on the same link and to detect duplicate addresses. Each IPv6 unicast or anycast address has a corresponding solicited-node address. The format of a solicited-node multicast address is: FF02:0:0:0:0:1:FFXX:XXXX where FF02:0:0:0:0:1:FF is fixed and consists of 104 bits, and XX:XXXX is the last 24 bits of an IPv6 unicast address or anycast address.

EUI-64 address-based interface identifiers

An interface identifier is 64 bits and uniquely identifies an interface on a link.

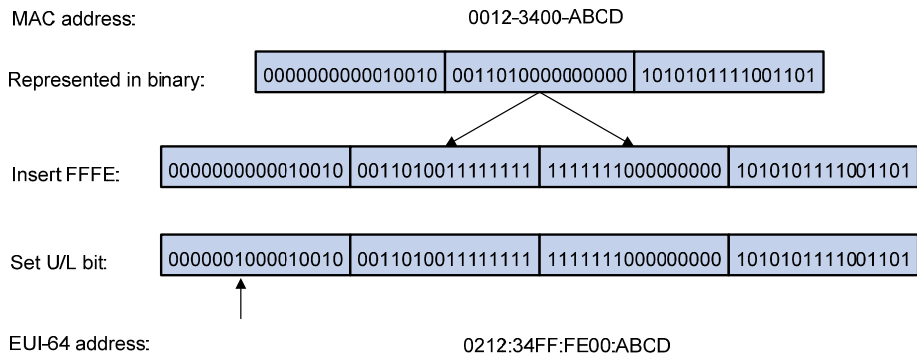
Interfaces generate EUI-64 address-based interface identifiers differently.

- On an IEEE 802 interface (such as a VLAN interface)

The interface identifier is derived from the link-layer address (typically a MAC address) of the interface. The MAC address is 48 bits long. To obtain an EUI-64 address-based interface identifier, you must insert the hexadecimal number FFFE (16 bits of 1111111111111110) into the MAC address (behind the 24th high-order bit), and set the universal/local (U/L) bit (which is the seventh high-order bit) to 1, to make sure that the obtained EUI-64 address-based interface identifier is globally unique.

[Figure 52](#) shows how an EUI-64 address-based interface identifier is generated from a MAC address.

Figure 52 Converting a MAC address into an EUI-64 address-based interface identifier



- On a tunnel interface
The lower 32 bits of the EUI-64 address-based interface identifier are the source IPv4 address of the tunnel interface. The higher 32 bits of the EUI-64 address-based interface identifier of an ISATAP tunnel interface are 0000:5EFE, whereas those of other tunnel interfaces are all zeros. For more information about tunnels, see "Configuring tunneling."
- On an interface of another type
The EUI-64 address-based interface identifier is generated randomly by the device.

IPv6 neighbor discovery protocol

The IPv6 Neighbor Discovery (ND) protocol uses five types of ICMPv6 messages to implement the following functions:

- Address resolution
- Neighbor reachability detection
- Duplicate address detection
- Router/prefix discovery and address autoconfiguration
- Redirection

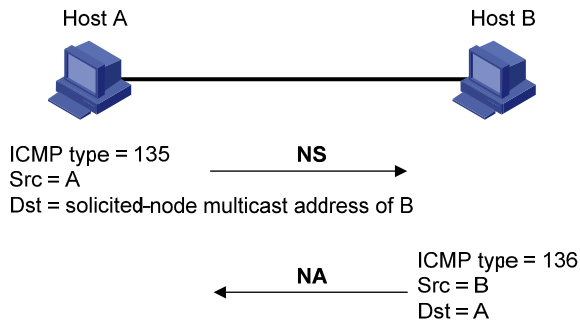
Table 8 ICMPv6 messages used by ND

ICMPv6 message	Type	Function
Neighbor Solicitation (NS) message	135	Acquires the link-layer address of a neighbor.
		Verifies whether a neighbor is reachable.
		Detects duplicate addresses.
Neighbor Advertisement (NA) message	136	Responds to an NS message.
		Notifies the neighboring nodes of link layer changes.
Router Solicitation (RS) message	133	Requests an address prefix and other configuration information for autoconfiguration after startup.
Router Advertisement (RA) message	134	Responds to an RS message.
		Advertises information such as the Prefix Information options and flag bits.
Redirect message	137	Informs the source host of a better next hop on the path to a particular destination when certain conditions are satisfied.

Address resolution

This function is similar to the ARP function in IPv4. An IPv6 node acquires the link-layer addresses of neighboring nodes on the same link through NS and NA message exchanges. Figure 53 shows how Host A acquires the link-layer address of Host B on a single link.

Figure 53 Address resolution



The address resolution operates in the following steps:

1. Host A multicasts an NS message. The source address of the NS message is the IPv6 address of the sending interface of Host A and the destination address is the solicited-node multicast address of Host B. The NS message contains the link-layer address of Host A.
2. After receiving the NS message, Host B determines whether the destination address of the packet is its solicited-node multicast address. If yes, Host B learns the link-layer address of Host A, and then unicasts an NA message containing its link-layer address.
3. Host A acquires the link-layer address of Host B from the NA message.

Neighbor reachability detection

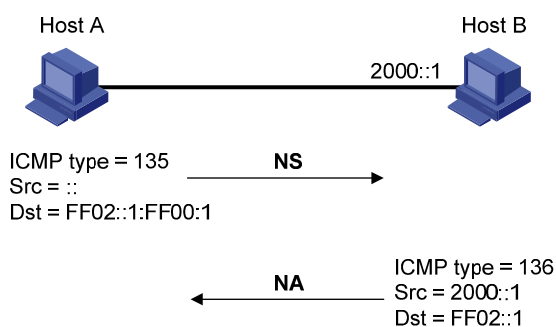
After Host A acquires the link-layer address of its neighbor Host B, Host A can use NS and NA messages to check whether Host B is reachable.

1. Host A sends an NS message whose destination address is the IPv6 address of Host B.
2. If Host A receives an NA message from Host B, Host A decides that Host B is reachable. Otherwise, Host B is unreachable.

Duplicate address detection

After Host A acquires an IPv6 address, it performs Duplicate Address Detection (DAD) to check whether the address is being used by any other node (similar to the gratuitous ARP function in IPv4). DAD is accomplished through NS and NA message exchanges. Figure 54 shows the DAD process.

Figure 54 Duplicate address detection



1. Host A sends an NS message whose source address is the unspecified address and whose destination address is the corresponding solicited-node multicast address of the IPv6 address to be detected. The NS message contains the IPv6 address.
2. If Host B uses this IPv6 address, Host B returns an NA message. The NA message contains the IPv6 address of Host B.
3. Host A learns that the IPv6 address is being used by Host B after receiving the NA message from Host B. If receiving no NA message, Host A decides that the IPv6 address is not in use and uses this address.

Router/prefix discovery and address autoconfiguration

Router/prefix discovery enables a node to locate the neighboring routers and to learn from the received RA message configuration parameters such as the prefix of the network where the node is located.

Stateless address autoconfiguration enables a node to generate an IPv6 address automatically according to the information obtained through router/prefix discovery.

Router/prefix discovery is implemented through RS and RA messages in the following steps:

1. At startup, a node sends an RS message to request the address prefix and other configuration information for autoconfiguration.
2. A router returns an RA message containing information such as Prefix Information options. (The router also periodically sends an RA message. In addition to an address prefix, the Prefix Information option also contains the preferred lifetime and valid lifetime of the address prefix. Nodes update the preferred lifetime and valid lifetime accordingly through periodic RA messages.)
3. The node automatically generates an IPv6 address and other configuration information according to the address prefix and other configuration parameters in the RA message. (The automatically generated address is applicable within the valid lifetime and is removed when the valid lifetime expires.)

Redirection

A newly started host may contain only a default route to the gateway in its routing table. When certain conditions are satisfied, the gateway sends an ICMPv6 Redirect message to the source host, so the host can select a better next hop to forward packets (similar to the ICMP redirection function in IPv4).

The gateway sends an ICMPv6 Redirect message when the following conditions are satisfied.

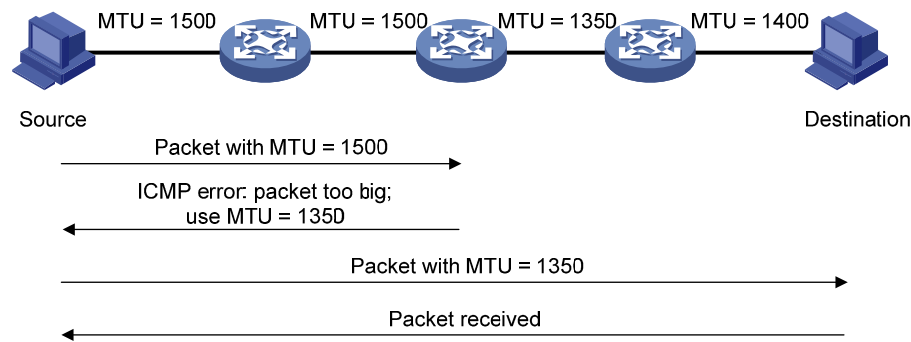
- The receiving interface is the forwarding interface.
- The selected route itself is not created or modified by an ICMPv6 Redirect message.
- The selected route is not the default route.
- The IPv6 packet to be forwarded does not contain any routing header.

IPv6 path MTU discovery

The links that a packet passes from a source to a destination may have different MTUs. In IPv6, when the packet size exceeds the path MTU of a link, the packet is fragmented at the source end of the link to reduce the processing pressure on intermediate devices and to use network resources effectively.

The path MTU discovery mechanism is designed to find the minimum MTU of all links in the path between a source and a destination. [Figure 55](#) shows how a source host discovers the path MTU to a destination host.

Figure 55 Path MTU discovery process



1. The source host compares its MTU with the packet to be sent, performs necessary fragmentation, and sends the resulting packet to the destination host.
2. If the MTU supported by a forwarding interface is smaller than the packet, the device discards the packet and returns an ICMPv6 error packet containing the interface MTU to the source host.
3. After receiving the ICMPv6 error packet, the source host uses the returned MTU to limit the packet size, performs fragmentation, and sends the resulting packet to the destination host.
4. Step 2 and step 3 are repeated until the destination host receives the packet. In this way, the source host decides the minimum MTU of all links in the path to the destination host.

IPv6 transition technologies

Before IPv6 dominates the Internet, high-efficient and seamless IPv6 transition technologies are needed to enable communication between IPv4 and IPv6 networks. Several IPv6 transition technologies can be used in different environments and periods, such as dual stack (RFC 2893) and tunneling (RFC 2893).

Dual stack

Dual stack is the most direct transition approach. A network node that supports both IPv4 and IPv6 is a dual stack node. A dual stack node configured with an IPv4 address and an IPv6 address can forward both IPv4 and IPv6 packets. For an upper layer application that supports both IPv4 and IPv6, either TCP or UDP can be selected at the transport layer, whereas the IPv6 stack is preferred at the network layer. Dual stack is suitable for communication between IPv4 nodes or between IPv6 nodes. It is the basis of all transition technologies. However, it does not solve the IPv4 address depletion issue because each dual stack node must have a globally unique IP address.

Tunneling

Tunneling is an encapsulation technology that utilizes one network protocol to encapsulate packets of another network protocol and transfer them over the network. For more information about tunneling, see "Configuring tunneling."

Protocols and standards

Protocols and standards related to IPv6 include:

- RFC 1881, *IPv6 Address Allocation Management*
- RFC 1887, *An Architecture for IPv6 Unicast Address Allocation*
- RFC 1981, *Path MTU Discovery for IP version 6*
- RFC 2375, *IPv6 Multicast Address Assignments*

- RFC 2460, *Internet Protocol, Version 6 (IPv6) Specification*
- RFC 2461, *Neighbor Discovery for IP Version 6 (IPv6)*
- RFC 2462, *IPv6 Stateless Address Autoconfiguration*
- RFC 2463, *Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification*
- RFC 2464, *Transmission of IPv6 Packets over Ethernet Networks*
- RFC 2526, *Reserved IPv6 Subnet Anycast Addresses*
- RFC 2894, *Router Renumbering for IPv6*
- RFC 3307, *Allocation Guidelines for IPv6 Multicast Addresses*
- RFC 3513, *Internet Protocol Version 6 (IPv6) Addressing Architecture*

IPv6 basics configuration task list

Task	Remarks	
Configuring basic IPv6 functions	Enabling IPv6	Required
	Configuring an IPv6 global unicast address	Required to configure one
	Configuring an IPv6 link-local address	
	Configure an IPv6 anycast address	
Configuring IPv6 ND	Configuring a static neighbor entry	Optional
	Configuring the maximum number of neighbors dynamically learned	Optional
	Setting the age timer for ND entries in stale state	Optional
	Configuring parameters related to RA messages	Optional
	Configuring the maximum number of attempts to send an NS message for DAD	Optional
	Configuring ND snooping	Optional
	Enabling ND proxy	Optional
Configuring path MTU discovery	Configuring a static path MTU for a specified IPv6 address	Optional
	Configuring the aging time for dynamic path MTUs	Optional
Configuring IPv6 TCP properties		Optional
Configuring ICMPv6 packet sending	Configuring the maximum ICMPv6 error packets sent in an interval	Optional
	Enabling replying to multicast echo requests	Optional
	Enabling sending of ICMPv6 time exceeded messages	Optional
	Enabling sending of ICMPv6 destination unreachable messages	Optional

Configuring basic IPv6 functions

Enabling IPv6

Enable IPv6 before you perform any IPv6-related configuration. Without IPv6 enabled, an interface cannot forward IPv6 packets even if it has an IPv6 address configured.

To enable IPv6:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IPv6.	ipv6	Disabled by default

Configuring an IPv6 global unicast address

Configure an IPv6 global unicast address by using the following options:

- **EUI-64 IPv6 addressing**—The IPv6 address prefix of an interface is manually configured, and the interface identifier is generated automatically by the interface.
- **Manual configuration**—The IPv6 global unicast address is configured manually.
- **Stateless address autoconfiguration**—The IPv6 global unicast address is generated automatically based on the address prefix information contained in the RA message.

Follow these guidelines when you configure an IPv6 global unicast address:

- You can configure multiple IPv6 global unicast addresses with different prefixes on an interface.
- A manually configured global unicast address takes precedence over an automatically generated one. If a global unicast address has been automatically generated on an interface when you manually configure another one with the same address prefix, the latter overwrites the previous. The overwritten automatic global unicast address will not be restored even if the manual one is removed. Instead, a new global unicast address will be automatically generated based on the address prefix information in the RA message that the interface receives at the next time.

EUI-64 IPv6 addressing

To configure an interface to generate an EUI-64 IPv6 address:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the interface to generate an EUI-64 IPv6 address.	ipv6 address <i>ipv6-address/prefix-length</i> eui-64	By default, no IPv6 global unicast address is configured on an interface.

Manual configuration

To specify an IPv6 address manually for an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure an IPv6 address manually.	ipv6 address { <i>ipv6-address</i> <i>prefix-length</i> <i>ipv6-address/prefix-length</i> }	By default, no IPv6 global unicast address is configured on an interface.

Stateless address autoconfiguration

To configure an interface to generate an IPv6 address by using stateless address autoconfiguration:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure an IPv6 address to be generated through stateless address autoconfiguration.	ipv6 address auto	By default, no IPv6 global unicast address is configured on an interface.

NOTE:

Using the **undo ipv6 address auto** command on an interface removes all IPv6 global unicast addresses automatically generated on the interface.

With stateless address autoconfiguration enabled on an interface, the device automatically generates an IPv6 global unicast address by using the address prefix information in the received RA message and the interface ID. On an IEEE 802 interface (such as a VLAN interface), the interface ID is generated based on the MAC address of the interface, and is globally unique. As a result, the interface ID portion of the IPv6 global address remains unchanged and exposes the sender. An attacker can further exploit communication details such as the communication peer and time.

To fix the vulnerability, configure the temporary address function that enables the system to generate and use temporary IPv6 addresses with different interface ID portions on an interface. With this function configured on an IEEE 802 interface, the system can generate two addresses, public IPv6 address and temporary IPv6 address.

- **Public IPv6 address**—Comprises an address prefix provided by the RA message, and a fixed interface ID generated based on the MAC address of the interface.
- **Temporary IPv6 address**—Comprises an address prefix provided by the RA message, and a random interface ID generated through MD5.

Before sending a packet, the system preferably uses the temporary IPv6 address of the sending interface as the source address of the packet to be sent. When this temporary IPv6 address expires, the system removes it and generates a new one. This enables the system to send packets with different source addresses through the same interface. If the temporary IPv6 address cannot be used because of a DAD conflict, the public IPv6 address is used.

The preferred lifetime and valid lifetime for temporary IPv6 addresses are specified as follows:

- The preferred lifetime of a temporary IPv6 address takes the value of the smaller of the following values:
 - The preferred lifetime of the address prefix in the RA message.

- The preferred lifetime configured for temporary IPv6 addresses minus DESYNC_FACTOR (which is a random number ranging 0 to 600, in seconds).
- The valid lifetime of a temporary IPv6 address takes the value of the smaller of the following values:
 - The valid lifetime of the address prefix.
 - The valid lifetime configured for temporary IPv6 addresses.

To configure the temporary address function:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the system to generate and preferably use the temporary IPv6 address of the sending interface as the source address of the packet to be sent.	ipv6 prefer temporary-address [<i>valid-lifetime preferred-lifetime</i>]	By default, the system does not generate or use a temporary IPv6 address.

You must also enable stateless address autoconfiguration on an interface if you need temporary IPv6 addresses to be generated on that interface. Temporary IPv6 addresses do not override public IPv6 addresses. Therefore, an interface may have multiple IPv6 addresses with the same address prefix but different interface ID portions.

If the public IPv6 address fails to be generated on an interface because of a prefix conflict or other reasons, no temporary IPv6 address will be generated on the interface.

Configuring an IPv6 link-local address

IPv6 link-local addresses can be configured in either of the following ways:

- **Automatic generation**—The device automatically generates a link-local address for an interface according to the link-local address prefix (FE80::/10) and the link-layer address of the interface.
- **Manual assignment**—IPv6 link-local addresses can be assigned manually.

An interface can have only one link-local address. To avoid link-local address conflicts, use the automatic generation method.

Manual assignment takes precedence over automatic generation.

- If you first use automatic generation and then manual assignment, the manually assigned link-local address will overwrite the automatically generated one.
- If you first use manual assignment and then automatic generation, the automatically generated link-local address will not take effect and the link-local address is still the manually assigned one. If you delete the manually assigned address, the automatically generated link-local address is validated.

To configure automatic generation of an IPv6 link-local address for an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A

Step	Command	Remarks
3.	Configure the interface to automatically generate an IPv6 link-local address. ipv6 address auto link-local	Optional. By default, no link-local address is configured on an interface. After an IPv6 global unicast address is configured on the interface, a link-local address is generated automatically.

To configure an IPv6 link-local address manually:

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Enter interface view. interface <i>interface-type</i> <i>interface-number</i>	N/A
3.	Configure an IPv6 link-local address manually. ipv6 address <i>ipv6-address</i> link-local	Optional. By default, no link-local address is configured on an interface. After an IPv6 global unicast address is configured on the interface, a link-local address is generated automatically.

After an IPv6 global unicast address is configured for an interface, a link-local address is generated automatically.

- The automatically generated link-local address is the same as the one generated by using the **ipv6 address auto link-local** command.
- If a link-local address is manually assigned to an interface, this manual link-local address takes effect. If the manually assigned link-local address is removed, the automatically generated link-local address takes effect.

The **undo ipv6 address auto link-local** command can only remove the link-local addresses generated through the **ipv6 address auto link-local** command.

- If an IPv6 global unicast address is already configured for an interface, the interface still has a link-local address because the system automatically generates one for the interface.
- If no IPv6 global unicast address is configured, the interface has no link-local address.

Configure an IPv6 anycast address

To configure an IPv6 anycast address for an interface:

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Enter interface view. interface <i>interface-type</i> <i>interface-number</i>	N/A
3.	Configure an IPv6 anycast address. ipv6 address <i>ipv6-address/prefix-length</i> anycast	Optional. By default, no IPv6 anycast address is configured on an interface.

Configuring IPv6 ND

Configuring a static neighbor entry

The IPv6 address of a neighboring node can be resolved into a link-layer address dynamically through NS and NA messages or through a manually configured static neighbor entry.

The device uniquely identifies a static neighbor entry by the neighbor's IPv6 address and the local Layer 3 interface number. You can configure a static neighbor entry by using either of the following methods:

- **Method 1**—Associate a neighbor IPv6 address and link-layer address with the Layer 3 interface of the local node.
- **Method 2**—Associate a neighbor IPv6 address and link-layer address with a port in a VLAN containing the local node.

You can use either of the previous configuration methods to configure a static neighbor entry for a VLAN interface.

- After a static neighbor entry is configured by using the first method, the device must resolve the corresponding Layer 2 port information of the VLAN interface.
- If you use the second method, make sure that the corresponding VLAN interface exists and that the Layer 2 port specified by *port-type port-number* belongs to the VLAN specified by *vlan-id*. After a static neighbor entry is configured, the device associates the VLAN interface with the IPv6 address to identify the static neighbor entry uniquely.

To configure a static neighbor entry:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure a static neighbor entry.	ipv6 neighbor <i>ipv6-address</i> <i>mac-address</i> { <i>vlan-id port-type</i> <i>port-number</i> interface <i>interface-type interface-number</i> } [vpn-instance <i>vpn-instance-name</i>]	Only HP 5500 EI Switch Series supports the vpn-instance <i>vpn-instance-name</i> option.

Configuring the maximum number of neighbors dynamically learned

The device can dynamically acquire the link-layer address of a neighboring node through NS and NA messages and add it into the neighbor table. A large table can reduce the forwarding performance of the device. You can restrict the size of the neighbor table by setting the maximum number of neighbors that an interface can dynamically learn. When the number of dynamically learned neighbors reaches the threshold, the interface will stop learning neighbor information.

To configure the maximum number of neighbors dynamically learned:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the maximum number of neighbors dynamically learned by an interface.	ipv6 neighbors max-learning-num <i>number</i>	Optional. By default, a Layer 2 interface does not limit the number of neighbors dynamically learned. A Layer 3 interface can learn up to 1024 neighbors dynamically for HP 5500 SI Switch Series, and up to 4096 neighbors dynamically for HP 5500 EI Switch Series.

Setting the age timer for ND entries in stale state

ND entries in stale state have an age timer. If an ND entry in stale state is not refreshed before the timer expires, it transits to the delay state. If it is still not refreshed in five seconds, the ND entry transits to the probe state, and the device sends an NS message for detection. If no response is received, the device removes the ND entry.

To set the age timer for ND entries in stale state:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the age timer for ND entries in stale state.	ipv6 neighbor stale-aging <i>aging-time</i>	Optional. Four hours by default.

Configuring parameters related to RA messages

You can enable an interface to send RA messages, and configure the interval for sending RA messages and parameters in RA messages. After receiving an RA message, a host can use these parameters to perform corresponding operations. [Table 9](#) lists and describes the configurable parameters in an RA message.

The maximum interval for sending RA messages should be less than (or equal to) the router lifetime in RA messages, so the router can be updated through an RA message before expiration.

The values of the NS retransmission timer and the reachable time configured for an interface are sent to hosts via RA messages. Furthermore, this interface sends NS messages at the interval of the NS retransmission timer and considers a neighbor reachable within the reachable time.

Table 9 Parameters in an RA message and their descriptions

Parameters	Description
Cur Hop Limit	When sending an IPv6 packet, a host uses the value to fill the Hop Limit field in IPv6 headers. The value is also filled into the Hop Limit field in the response packet of a device.
Prefix Information options	After receiving the prefix information, the hosts on the same link can perform stateless autoconfiguration.

Parameters	Description
MTU	Make sure that all nodes on a link use the same MTU value.
M flag	Determines whether hosts use the stateful autoconfiguration to acquire IPv6 addresses. If the M flag is set to 1, hosts use the stateful autoconfiguration (for example, through a DHCP server) to acquire IPv6 addresses. Otherwise, hosts use the stateless autoconfiguration to acquire IPv6 addresses and generate IPv6 addresses according to their own link-layer addresses and the obtained prefix information.
O flag	Determines whether hosts use stateful autoconfiguration to acquire other configuration information. If the O flag is set to 1, hosts use stateful autoconfiguration (for example, through a DHCP server) to acquire other configuration information. Otherwise, hosts use stateless autoconfiguration to acquire other configuration information.
Router Lifetime	Tells the receiving hosts how long the advertising device can live
Retrans Timer	If the device fails to receive a response message within the specified time after sending an NS message, it will retransmit the NS message.
Reachable Time	If the neighbor reachability detection shows that a neighbor is reachable, the device considers the neighbor reachable within the specified reachable time. If the device must send a packet to the neighbor after the specified reachable time expires, the device will reconfirm whether the neighbor is reachable.

To allow sending of RA messages:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Disable RA message suppression.	undo ipv6 nd ra halt	By default, RA messages are suppressed.
4. Configure the maximum and minimum intervals for sending RA messages.	ipv6 nd ra interval <i>max-interval-value</i> <i>min-interval-value</i>	Optional. By default, the maximum interval for sending RA messages is 600 seconds, and the minimum interval is 200 seconds. The device sends RA messages at random intervals between the maximum interval and the minimum interval. The minimum interval should be less than or equal to 0.75 times the maximum interval.

To configure parameters related to RA messages:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the hop limit.	ipv6 nd hop-limit <i>value</i>	Optional. 64 by default.

Step	Command	Remarks
3. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Configure the prefix information in RA messages.	ipv6 nd ra prefix { <i>ipv6-prefix</i> <i>prefix-length</i> <i>ipv6-prefix/prefix-length</i> } <i>valid-lifetime preferred-lifetime</i> [no-autoconfig off-link] *	Optional. By default, no prefix information is configured for RA messages, and the IPv6 address of the interface sending RA messages is used as the prefix information with valid lifetime 2592000 seconds (30 days) and preferred lifetime 604800 seconds (seven days).
5. Turn off the MTU option in RA messages.	ipv6 nd ra no-advlinkmtu	Optional. By default, RA messages contain the MTU option.
6. Set the M flag bit to 1.	ipv6 nd autoconfig managed-address-flag	Optional. By default, the M flag bit is set to 0 and hosts acquire IPv6 addresses through stateless autoconfiguration.
7. Set the O flag bit to 1.	ipv6 nd autoconfig other-flag	Optional. By default, the O flag bit is set to 0 and hosts acquire other configuration information through stateless autoconfiguration.
8. Configure the router lifetime in RA messages.	ipv6 nd ra router-lifetime <i>value</i>	Optional. 1800 seconds by default.
9. Set the NS retransmission timer.	ipv6 nd ns retrans-timer <i>value</i>	Optional. By default, the local interface sends NS messages at 1000 millisecond intervals, and the value of the Retrans Timer field in RA messages sent by the local interface is 0. The interval for retransmitting an NS message is determined by the receiving device.
10. Set the reachable time.	ipv6 nd nud reachable-time <i>value</i>	Optional. By default, the neighbor reachable time on the local interface is 30000 milliseconds, and the value of the Reachable Time field in the RA messages sent by the local interface is 0. The neighbor reachable time is determined by the receiving device.

Configuring the maximum number of attempts to send an NS message for DAD

An interface sends an NS message for DAD after acquiring an IPv6 address. If the interface does not receive a response within a specified time (determined by the **ipv6 nd ns retrans-timer** command), it

continues to send an NS message. If the interface still does not receive a response after the number of sent attempts reaches the threshold (specified with the **ipv6 nd dad attempts** command), the acquired address is considered usable.

To configure the attempts to send an NS message for DAD:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the number of attempts to send an NS message for DAD.	ipv6 nd dad attempts <i>value</i>	Optional. 1 by default. When the <i>value</i> argument is set to 0, DAD is disabled.

Configuring ND snooping

Introduction

The ND snooping feature is used in Layer 2 switching networks. It creates ND snooping entries using DAD NS messages.

ND snooping entries are used to do the following:

- Cooperate with the ND detection function. For more information about ND detection, see *Security Configuration Guide*.
- Cooperate with the IP Source Guard function. For more information about IP source guard, see *Security Configuration Guide*.
- Work in all SAVI scenarios. For more information about SAVI, see *Security Configuration Guide*.

After you enable ND snooping on a VLAN of a device, ND packets received by the interfaces of the VLAN are redirected to the CPU. When ND snooping is enabled globally, the CPU uses the ND packets to create or update ND snooping entries comprising source IPv6 address, source MAC address, receiving VLAN, and receiving port information.

The following items describe how an ND snooping entry is created, updated, and aged out.

1. Create an ND snooping entry

The device only uses received DAD NS messages to create ND snooping entries.

2. Update an ND snooping entry

Upon receiving an ND packet, the device searches the ND snooping table for an entry containing the source IPv6 address of the packet. If the entry was refreshed within one second, the device does not update the entry. If the entry is not refreshed for more than one second, the device matches the MAC address of the ND packet and the receiving port against that in the entry.

- If both of them match those in the entry, the device updates the aging time of the ND snooping entry.
- If neither of them matches the entry and the received packet is a DAD NS message, the message is ignored.
- If neither of them matches the entry and the received packet is not a DAD NS message, the device performs active acknowledgement.

The active acknowledgement is performed in the following steps.

- The device checks the validity of the existing ND snooping entry. The device sends out a DAD NS message including the IPv6 address of the ND snooping entry. If a corresponding NA message (whose source IPv6 address, source MAC address, receiving port, and source VLAN are consistent with those of the existing entry) is received, the device updates the aging time of the existing entry. If no corresponding NA message is received within one second after the DAD NS message is sent, the device starts to check the validity of the received ND packet.
 - To check the validity of the received ND packet (packet A for example), the device sends out a DAD NS message including the source IPv6 address of packet A. If a corresponding NA message (whose source IPv6 address, source MAC address, receiving port, and source VLAN are consistent with those of packet A) is received, the device updates the aging time of the entry. If no corresponding NA message is received within one second after the DAD NS message is sent, the device does not update the entry.
- 3. Age out an ND snooping entry**
- An ND snooping entry is aged out after 25 minutes. If an ND snooping entry is not updated within 15 minutes, the device performs active acknowledgement.
- The device sends out a DAD NS message including the IPv6 address of the ND snooping.
- If a corresponding NA message is received (the source IPv6 address, source MAC address, receiving port, and source VLAN are consistent with those of the existing entry), the device updates the aging time of the existing entry.
 - If no corresponding NA message is received within one second after the DAD NS message is sent out, the device removes the entry when the timer expires.

Configuration procedure

To configure ND snooping:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure ND snooping.	<ul style="list-style-type: none"> • Enable ND snooping based on global unicast addresses (the devices use DAD NS messages containing global unicast addresses to create ND snooping entries): ipv6 nd snooping enable global • Enable ND snooping based on link local addresses (the devices use DAD NS messages containing link local addresses to create ND snooping entries): ipv6 nd snooping enable link-local 	<p>Use either approach.</p> <p>By default, ND snooping is disabled.</p>
3. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
4. Enable ND snooping.	ipv6 nd snooping enable	Disabled by default.
5. Return to system view.	quit	N/A
6. Enter Layer 2 Ethernet port view/Layer 2 aggregate interface view.	interface <i>interface-type interface-number</i>	N/A

Step	Command	Remarks
7. Configure the maximum number of ND snooping entries the interface can learn.	<code>ipv6 nd snooping max-learning-num number</code>	Optional. By default, the number of ND snooping entries an interface can learn is unlimited.
8. Configure the interface as an uplink interface and disable it from learning ND snooping entries.	<code>ipv6 nd snooping uplink</code>	Optional. By default, when ND snooping is enabled on the device, an interface is allowed to learn ND snooping entries.

Enabling ND proxy

ND proxy supports the NS and NA messages only.

Introduction

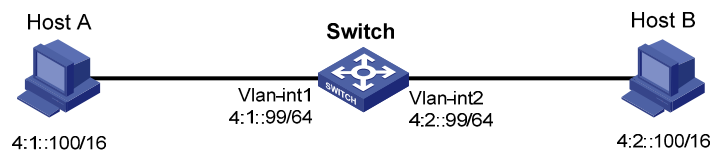
If a host sends an NS message requesting the hardware address of another host that is isolated from the sending host at Layer 2, the device between the hosts must be able to forward the NS message to allow Layer 3 communication between the two hosts. This process is achieved by ND proxy.

Depending on application scenarios, ND proxy falls into common ND proxy and local ND proxy.

- Common ND proxy

As shown in [Figure 56](#), VLAN-interface 1 with IPv6 address 4:1::99/64 and VLAN-interface 2 with IPv6 address 4:2::99/64 belong to different subnets. Host A and Host B reside on the same network but in different broadcast domains.

Figure 56 Application environment of common ND proxy



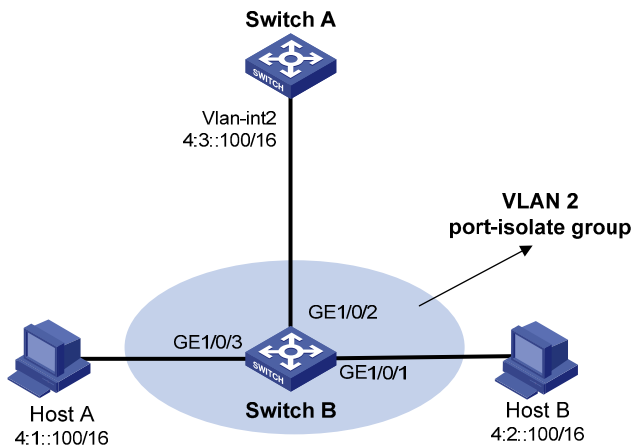
Because Host A's IPv6 address is on the same subnet as Host B's, Host A directly sends an NS message to obtain Host B's MAC address. However, Host B cannot receive the NS message because they belong to different broadcast domains.

To solve this problem, enable common ND proxy on VLAN-interface 1 and VLAN-interface 2 of the switch. The switch finds the matching forwarding entry according to the destination IPv6 address of the NS message and sends the message through the output interface of that entry. Upon receiving the NS message, Host B sends an NA message to the switch, which forwards it to Host A.

- Local ND proxy

As shown in [Figure 57](#), both Host A and Host B belong to VLAN 2, but they connect to GigabitEthernet 1/0/3 and GigabitEthernet 1/0/1 respectively, which are isolated at Layer 2.

Figure 57 Application environment of local ND proxy



Because Host A's IPv6 address is on the same subnet as Host B's, Host A directly sends an NS message to obtain Host B's MAC address. However, Host B cannot receive the NS message because they are isolated at Layer 2.

To solve this problem, enable local ND proxy on VLAN-interface 2 of the switch A so that the switch A can forward messages between Host A and Host B.

Local ND proxy implements Layer 3 communication for two hosts in the following cases:

- The two hosts must connect to different isolated Layer 2 ports of a VLAN.
- If super VLAN is used, the two hosts must belong to different sub VLANs.
- If isolate-user-VLAN is used, the two hosts must belong to different secondary VLANs.

Configuration procedure

You can enable common ND proxy and local ND proxy in VLAN interface view, Layer 3 Ethernet port view.

To enable common ND proxy:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable common ND proxy.	proxy-nd enable	Disabled by default

To enable local ND proxy:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable local ND proxy.	local-proxy-nd enable	Optional. Disabled by default.

Configuring path MTU discovery

Configuring a static path MTU for a specified IPv6 address

You can configure a static path MTU for a specified destination IPv6 address. When a source host sends a packet through an interface, it compares the interface MTU with the static path MTU of the specified destination IPv6 address. If the packet size is larger than the smaller one of the two values, the host fragments the packet according to the smaller value.

To configure a static path MTU for a specified IPv6 address:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure a static path MTU for a specified IPv6 address.	ipv6 pathmtu [vpn-instance <i>vpn-instance-name</i>] <i>ipv6-address</i> [<i>value</i>]	Not configured by default Only HP 5500 EI Switch Series supports the vpn-instance <i>vpn-instance-name</i> option

Configuring the aging time for dynamic path MTUs

After the path MTU from a source host to a destination host is dynamically determined (see "[IPv6 path MTU discovery](#)"), the source host sends subsequent packets to the destination host based on this MTU. After the aging time expires, the dynamic path MTU is removed and the source host re-determines a dynamic path MTU through the path MTU mechanism.

The aging time is invalid for a static path MTU.

To configure the aging time for dynamic path MTUs:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the aging time for dynamic path MTUs.	ipv6 pathmtu age <i>age-time</i>	Optional. 10 minutes by default.

Configuring IPv6 TCP properties

You can configure the following IPv6 TCP properties:

- **synwait timer**—When a SYN packet is sent, the synwait timer is triggered. If no response packet is received before the synwait timer expires, the IPv6 TCP connection establishment fails.
- **finwait timer**—When the IPv6 TCP connection status is FIN_WAIT_2, the finwait timer is triggered. If no packet is received before the finwait timer expires, the IPv6 TCP connection is terminated. If a FIN packet is received, the IPv6 TCP connection status becomes TIME_WAIT. If non-FIN packets are received, the finwait timer is reset upon receipt of the last non-FIN packet and the connection is terminated after the finwait timer expires.
- **Size of the IPv6 TCP sending/receiving buffer**

To configure IPv6 TCP properties:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the synwait timer.	tcp ipv6 timer syn-timeout <i>wait-time</i>	Optional. 75 seconds by default.
3. Set the finwait timer.	tcp ipv6 timer fin-timeout <i>wait-time</i>	Optional. 675 seconds by default.
4. Set the size of the IPv6 TCP sending/receiving buffer.	tcp ipv6 window <i>size</i>	Optional. 8 KB by default.

Configuring ICMPv6 packet sending

Configuring the maximum ICMPv6 error packets sent in an interval

If too many ICMPv6 error packets are sent within a short time in a network, network congestion may occur. To avoid network congestion, you can control the maximum number of ICMPv6 error packets sent within a specified time by adopting the token bucket algorithm.

You can set the capacity of a token bucket to determine the number of tokens in the bucket. In addition, you can set the update interval of the token bucket, the interval for restoring the configured capacity. One token allows one ICMPv6 error packet to be sent. Each time an ICMPv6 error packet is sent, the number of tokens in a token bucket decreases by one. If the number of ICMPv6 error packets successively sent exceeds the capacity of the token bucket, the additional ICMPv6 error packets cannot be sent out until the capacity of the token bucket is restored.

To configure the capacity and update interval of the token bucket:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the capacity and update interval of the token bucket.	ipv6 icmp-error { bucket <i>bucket-size</i> ratelimit <i>interval</i> } *	Optional. By default, the capacity of a token bucket is 10 and the update interval is 100 milliseconds. A maximum of 10 ICMPv6 error packets can be sent within 100 milliseconds. The update interval "0" indicates that the number of ICMPv6 error packets sent is not restricted.

Enabling replying to multicast echo requests

If hosts are configured to answer multicast echo requests, an attacker can use this mechanism to attack a host. For example, if Host A (an attacker) sends an echo request with the source being Host B to a multicast address, all the hosts in the multicast group will send echo replies to Host B. To prevent such an attack, disable a device from answering multicast echo requests by default. In some application scenarios, however, you must enable the device to answer multicast echo requests.

To enable replying to multicast echo requests:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable replying to multicast echo requests.	ipv6 icmpv6 multicast-echo-reply enable	Not enabled by default

Enabling sending of ICMPv6 time exceeded messages

A device sends out an ICMPv6 Time Exceeded message in the following situations:

- If a received IPv6 packet's destination IP address is not a local address and its hop limit is 1, the device sends an ICMPv6 Hop Limit Exceeded message to the source.
- Upon receiving the first fragment of an IPv6 datagram with the destination IP address being the local address, the device starts a timer. If the timer expires before all the fragments arrive, an ICMPv6 Fragment Reassembly Timeout message is sent to the source.

If large quantities of malicious packets are received, the performance of a device degrades greatly because it must send back ICMP Time Exceeded messages. You can disable sending of ICMPv6 Time Exceeded messages.

To enable sending of ICMPv6 time exceeded messages:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable sending of ICMPv6 Time Exceeded messages.	ipv6 hoplimit-expires enable	Optional. Enabled by default.

Enabling sending of ICMPv6 destination unreachable messages

If the device fails to forward a received IPv6 packet because of one of the following reasons, it drops the packet and sends a corresponding ICMPv6 Destination Unreachable error message to the source.

- If no route is available for forwarding the packet, the device sends a "no route to destination" ICMPv6 error message to the source.
- If the device fails to forward the packet because of an administrative prohibition (such as a firewall filter or an ACL), the device sends the source a "destination network administratively prohibited" ICMPv6 error message.
- If the device fails to deliver the packet because the destination is beyond the scope of the source IPv6 address (for example, the source IPv6 address of the packet is a link-local address whereas the destination IPv6 address of the packet is a global unicast address), the device sends the source a "beyond scope of source address" ICMPv6 error message.
- If the device fails to resolve the corresponding link layer address of the destination IPv6 address, the device sends the source an "address unreachable" ICMPv6 error message.
- If the packet with the destination being local and transport layer protocol being UDP and the packet's destination port number does not match the running process, the device sends the source a "port unreachable" ICMPv6 error message.

If an attacker sends abnormal traffic that causes the device to generate ICMPv6 destination unreachable messages, end users may be affected. To prevent such attacks, you can disable the device from sending ICMPv6 destination unreachable messages.

To enable sending of ICMPv6 destination unreachable messages:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable sending of ICMPv6 destination unreachable messages.	ipv6 unreachable enable	Disabled by default

Displaying and maintaining IPv6 basics configuration

Task	Command	Remarks
Display the IPv6 FIB entries.	display ipv6 fib [vpn-instance <i>vpn-instance-name</i>] [acl6 <i>acl6-number</i> ipv6-prefix <i>ipv6-prefix-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view Only HP 5500 EI Switch Series supports the vpn-instance <i>vpn-instance-name</i> option.
Display the IPv6 FIB entry of a specified destination IPv6 address.	display ipv6 fib [vpn-instance <i>vpn-instance-name</i>] <i>ipv6-address</i> [<i>prefix-length</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view Only HP 5500 EI Switch Series supports the vpn-instance <i>vpn-instance-name</i> option.
Display the IPv6 information of the interface.	display ipv6 interface [<i>interface-type</i> [<i>interface-number</i>]] [brief] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display neighbor information.	display ipv6 neighbors { { <i>ipv6-address</i> all dynamic static } [<i>slot slot-number</i>] interface <i>interface-type interface-number</i> vlan <i>vlan-id</i> } [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view Only HP 5500 EI Switch Series supports the verbose keyword.
Display the total number of neighbor entries satisfying the specified conditions.	display ipv6 neighbors { { all dynamic static } [<i>slot slot-number</i>] interface <i>interface-type interface-number</i> vlan <i>vlan-id</i> } count [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the neighbor information of a specified VPN(only available on the HP 5500 EI).	display ipv6 neighbors vpn-instance <i>vpn-instance-name</i> [count] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display the IPv6 path MTU information.	display ipv6 pathmtu [vpn-instance <i>vpn-instance-name</i>] { <i>ipv6-address</i> all dynamic static } [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view Only HP 5500 EI Switch Series supports the vpn-instance <i>vpn-instance-name</i> option.
Display socket information.	display ipv6 socket [socket-type <i>socket-type</i>] [<i>task-id</i> <i>socket-id</i>] [slot <i>slot-number</i>] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display the statistics of IPv6 packets and ICMPv6 packets.	display ipv6 statistics [slot <i>slot-number</i>] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display the IPv6 TCP connection statistics.	display tcp ipv6 statistics [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display the IPv6 TCP connection status information.	display tcp ipv6 status [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display the IPv6 UDP connection statistics.	display udp ipv6 statistics [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display ND snooping entries.	display ipv6 nd snooping [<i>ipv6-address</i> vlan <i>vlan-id</i>] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Clear IPv6 neighbor information.	reset ipv6 neighbors { all dynamic interface <i>interface-type</i> <i>interface-number</i> slot <i>slot-number</i> static }	Available in user view
Clear the path MTU values.	reset ipv6 pathmtu { all static dynamic }	Available in user view
Clear the statistics of IPv6 and ICMPv6 packets.	reset ipv6 statistics [slot <i>slot-number</i>]	Available in user view
Clear all IPv6 TCP connection statistics.	reset tcp ipv6 statistics	Available in user view
Clear the statistics of all IPv6 UDP packets.	reset udp ipv6 statistics	Available in user view
Clear ND snooping entries.	reset ipv6 nd snooping [<i>ipv6-address</i> vlan <i>vlan-id</i>]	Available in user view

IPv6 basics configuration example

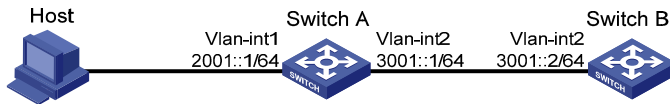
Network requirements

As shown in [Figure 58](#), a host, Switch A and Switch B are connected through Ethernet ports. Add the Ethernet ports into corresponding VLANs, configure IPv6 addresses for the VLAN interfaces and verify that they are connected.

- The global unicast addresses of VLAN-interface 1 and VLAN-interface 2 on Switch A are 2001::1/64 and 3001::1/64, respectively.
- The global unicast address of VLAN-interface 2 on Switch B is 3001::2/64, and a route to Host is available.

- IPv6 is enabled for the host to automatically obtain an IPv6 address through IPv6 ND, and a route to Switch B is available.

Figure 58 Network diagram



The VLAN interfaces have been created on the switch.

Configuration procedure

1. Configure Switch A:

Enable IPv6.

```
<SwitchA> system-view
[SwitchA] ipv6
```

Specify a global unicast address for VLAN-interface 2.

```
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ipv6 address 3001::1/64
[SwitchA-Vlan-interface2] quit
```

Specify a global unicast address for VLAN-interface 1, and allow it to advertise RA messages (no interface advertises RA messages by default).

```
[SwitchA] interface vlan-interface 1
[SwitchA-Vlan-interface1] ipv6 address 2001::1/64
[SwitchA-Vlan-interface1] undo ipv6 nd ra halt
[SwitchA-Vlan-interface1] quit
```

2. Configure Switch B:

Enable IPv6.

```
<SwitchB> system-view
[SwitchB] ipv6
```

Configure a global unicast address for VLAN-interface 2.

```
[SwitchB] interface vlan-interface 2
[SwitchB-Vlan-interface2] ipv6 address 3001::2/64
[SwitchB-Vlan-interface2] quit
```

Configure an IPv6 static route with destination IP address 2001::/64 and next hop address 3001::1.

```
[SwitchB] ipv6 route-static 2001:: 64 3001::1
```

3. Configure the host:

Enable IPv6 for Host to automatically obtain an IPv6 address through IPv6 ND.

Display the neighbor information of GigabitEthernet 1/0/2 on Switch A.

```
[SwitchA] display ipv6 neighbors interface GigabitEthernet 1/0/2
```

```

Type: S-Static      D-Dynamic
IPv6 Address          Link-layer      VID  Interface  State T Age
FE80::215:E9FF:FEA6:7D14  0015-e9a6-7d14  1    GE1/0/2    STALE D 1238
2001::15B:E0EA:3524:E791  0015-e9a6-7d14  1    GE1/0/2    STALE D 1248

```

The output shows that the IPv6 global unicast address that the host obtained is 2001::15B:E0EA:3524:E791.

Verifying the configuration

Display the IPv6 interface settings on Switch A. All of the IPv6 global unicast addresses configured on the interface are displayed.

```
[SwitchA] display ipv6 interface vlan-interface 2
Vlan-interface2 current state :UP
Line protocol current state :UP
IPv6 is enabled, link-local address is FE80::20F:E2FF:FE00:2
Global unicast address(es):
  3001::1, subnet is 3001::/64
Joined group address(es):
  FF02::1:FF00:0
  FF02::1:FF00:1
  FF02::1:FF00:2
  FF02::2
  FF02::1
MTU is 1500 bytes
ND DAD is enabled, number of DAD attempts: 1
ND reachable time is 30000 milliseconds
ND retransmit interval is 1000 milliseconds
Hosts use stateless autoconfig for addresses
IPv6 Packet statistics:
InReceives:                25829
InTooShorts:                0
InTruncatedPkts:           0
InHopLimitExceeds:         0
InBadHeaders:              0
InBadOptions:              0
ReasmReqds:                0
ReasmOKs:                  0
InFragDrops:               0
InFragTimeouts:           0
OutFragFails:              0
InUnknownProtos:          0
InDelivers:                47
OutRequests:               89
OutForwDatagrams:          48
InNoRoutes:                0
InTooBigErrors:            0
OutFragOKs:                0
OutFragCreates:            0
InMcastPkts:               6
InMcastNotMembers:         25747
OutMcastPkts:              48
InAddrErrors:              0
InDiscards:                0
OutDiscards:               0
```

```

[SwitchA] display ipv6 interface vlan-interface 1
Vlan-interfacel current state :UP
Line protocol current state :UP
IPv6 is enabled, link-local address is FE80::20F:E2FF:FE00:1C0
Global unicast address(es):
  2001::1, subnet is 2001::/64
Joined group address(es):
  FF02::1:FF00:0
  FF02::1:FF00:1
  FF02::1:FF00:1C0
  FF02::2
  FF02::1
MTU is 1500 bytes
ND DAD is enabled, number of DAD attempts: 1
ND reachable time is 30000 milliseconds
ND retransmit interval is 1000 milliseconds
ND advertised reachable time is 0 milliseconds
ND advertised retransmit interval is 0 milliseconds
ND router advertisements are sent every 600 seconds
ND router advertisements live for 1800 seconds
Hosts use stateless autoconfig for addresses
IPv6 Packet statistics:
InReceives:                272
InTooShorts:                0
InTruncatedPkts:           0
InHopLimitExceeds:         0
InBadHeaders:               0
InBadOptions:               0
ReasmReqds:                 0
ReasmOKs:                   0
InFragDrops:                0
InFragTimeouts:             0
OutFragFails:               0
InUnknownProtos:           0
InDelivers:                 159
OutRequests:                1012
OutForwDatagrams:           35
InNoRoutes:                 0
InTooBigErrors:             0
OutFragOKs:                 0
OutFragCreates:             0
InMcastPkts:                79
InMcastNotMembers:         65
OutMcastPkts:               938
InAddrErrors:               0
InDiscards:                 0
OutDiscards:                 0

```

Display the IPv6 interface settings on Switch B. All the IPv6 global unicast addresses configured on the interface are displayed.

```
[SwitchB] display ipv6 interface vlan-interface 2
Vlan-interface2 current state :UP
Line protocol current state :UP
IPv6 is enabled, link-local address is FE80::20F:E2FF:FE00:1234
  Global unicast address(es):
    3001::2, subnet is 3001::/64
  Joined group address(es):
    FF02::1:FF00:0
    FF02::1:FF00:2
    FF02::1:FF00:1234
    FF02::2
    FF02::1
  MTU is 1500 bytes
  ND DAD is enabled, number of DAD attempts: 1
  ND reachable time is 30000 milliseconds
  ND retransmit interval is 1000 milliseconds
  Hosts use stateless autoconfig for addresses
IPv6 Packet statistics:
  InReceives:                117
  InTooShorts:                0
  InTruncatedPkts:           0
  InHopLimitExceeds:         0
  InBadHeaders:               0
  InBadOptions:               0
  ReasmReqds:                  0
  ReasmOKs:                    0
  InFragDrops:                 0
  InFragTimeouts:             0
  OutFragFails:                0
  InUnknownProtos:           0
  InDelivers:                  117
  OutRequests:                 83
  OutForwDatagrams:           0
  InNoRoutes:                  0
  InTooBigErrors:              0
  OutFragOKs:                   0
  OutFragCreates:              0
  InMcastPkts:                 28
  InMcastNotMembers:          0
  OutMcastPkts:                 7
  InAddrErrors:                 0
  InDiscards:                   0
  OutDiscards:                   0
```

Ping Switch A and Switch B on the host, and ping Switch A and the host on Switch B to verify that they are connected.

❗ **IMPORTANT:**

When you ping a link-local address, you should use the **-i** parameter to specify an interface for the link-local address.

```
[SwitchB] ping ipv6 -c 1 3001::1
PING 3001::1 : 56 data bytes, press CTRL_C to break
  Reply from 3001::1
    bytes=56 Sequence=1 hop limit=64  time = 2 ms

--- 3001::1 ping statistics ---
  1 packet(s) transmitted
  1 packet(s) received
  0.00% packet loss
  round-trip min/avg/max = 2/2/2 ms
[SwitchB-Vlan-interface2] ping ipv6 -c 1 2001::15B:E0EA:3524:E791
PING 2001::15B:E0EA:3524:E791 : 56 data bytes, press CTRL_C to break
  Reply from 2001::15B:E0EA:3524:E791
    bytes=56 Sequence=1 hop limit=63  time = 3 ms

--- 2001::15B:E0EA:3524:E791 ping statistics ---
  1 packet(s) transmitted
  1 packet(s) received
  0.00% packet loss
  round-trip min/avg/max = 3/3/3 ms
```

The output shows that Switch B can ping Switch A and the host.

Troubleshooting IPv6 basics configuration

Symptom

The peer IPv6 address cannot be pinged.

Solution

1. Use the **display current-configuration** command in any view or the **display this** command in system view to verify that IPv6 is enabled.
2. Use the **display ipv6 interface** command in any view to verify that the IPv6 address of the interface is correct and the interface is up.
3. Use the **debugging ipv6 packet** command in user view to enable the debugging for IPv6 packets to help locate the cause.

DHCPv6 overview

Introduction to DHCPv6

The Dynamic Host Configuration Protocol for IPv6 (DHCPv6) was designed based on IPv6 addressing scheme and is used for assigning IPv6 prefixes, IPv6 addresses and other configuration parameters to hosts.

Compared with other IPv6 address allocation methods (such as manual configuration and stateless address autoconfiguration), DHCPv6 can:

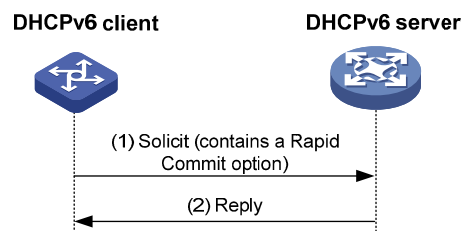
- Record addresses assigned to hosts and assign specific addresses to hosts, thus facilitating network management.
- Assign prefixes to devices, facilitating automatic configuration and management of the entire network.
- Assign other configuration parameters, such as DNS server addresses and domain names.

DHCPv6 address/prefix assignment

A process of DHCPv6 address/prefix assignment involves two or four messages. The following describe the detailed processes.

Rapid assignment involving two messages

Figure 59 Rapid assignment involving two messages



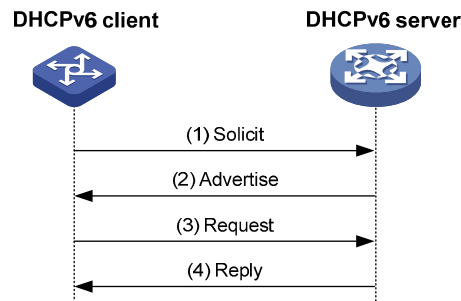
As shown in [Figure 59](#), the rapid assignment involving two messages operates in the following steps.

1. The DHCPv6 client sends out a Solicit message that contains a Rapid Commit option, requesting that rapid assignment of address/prefix and other configuration parameters should be preferred.
2. If the DHCPv6 server supports rapid assignment, it responds with a Reply message containing the assigned IPv6 address/prefix and other configuration parameters. If the DHCPv6 server does not support rapid assignment, [Assignment involving four messages](#) is implemented.

Assignment involving four messages

[Figure 60](#) shows the process of IPv6 address/prefix assignment involving four messages.

Figure 60 Assignment involving four messages



The assignment involving four messages operates in the following steps:

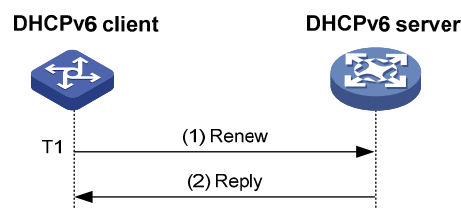
1. The DHCPv6 client sends out a Solicit message, requesting an IPv6 address/prefix and other configuration parameters.
2. If the Solicit message does not contain a Rapid Commit option, or if the DHCPv6 server does not support rapid assignment even though the Solicit message contains a Rapid Commit option, the DHCPv6 server responds with an Advertise message, informing the DHCPv6 client of the assignable address/prefix and other configuration parameters.
3. The DHCPv6 client may receive multiple Advertise messages offered by different DHCPv6 servers. It then selects an offer according to the receiving sequence and server priority, and sends a Request message to the selected server for the confirmation of assignment.
4. The DHCPv6 server sends a Reply message to the client, confirming that the address/prefix and other configuration parameters are assigned to the client.

Address/prefix lease renewal

The IPv6 address/prefix assigned by the DHCPv6 server has a lease time, which depends on the valid lifetime. When the valid lifetime of the IPv6 address/prefix expires, the DHCPv6 client cannot use the IPv6 address/prefix any longer. To continue using the IPv6 address/prefix, the DHCPv6 client has to renew the lease time.

As shown in [Figure 61](#), at T1, the DHCPv6 client unicasts a Renew message to the DHCPv6 server that assigned the IPv6 address/prefix to the DHCPv6 client. The recommended value of T1 is half the preferred lifetime. Then the DHCPv6 server responds with a Reply message, informing the client about whether or not the lease is renewed.

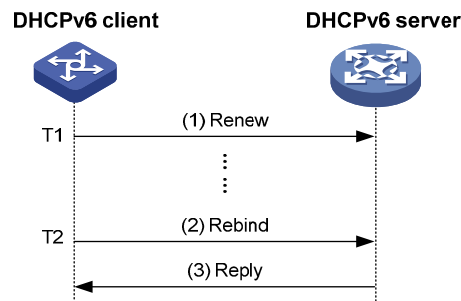
Figure 61 Using the Renew message for address/prefix lease renewal



As shown in [Figure 62](#), if the DHCPv6 client receives no response from the DHCPv6 server after sending out a Renew message at T1, it multicasts a Rebind message to all DHCPv6 servers at T2 (that is, when 80% preferred lifetime expires). Then the DHCPv6 server responds with a Reply message, informing the client about whether or not the lease is renewed.

If the DHCPv6 client receives no response from the DHCPv6 servers, the client stops using the address/prefix when the valid lifetime expires. For more information about the valid lifetime and the preferred lifetime, see "Configuring IPv6 basics."

Figure 62 Using the Rebind message for address/prefix lease renewal



Configuring stateless DHCPv6

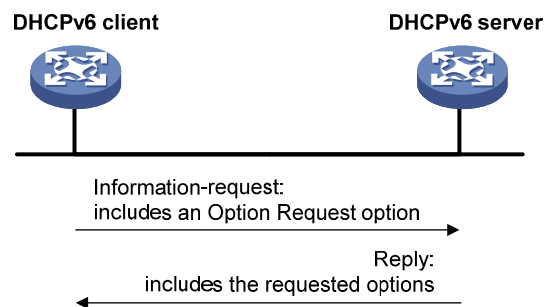
After obtaining an IPv6 address/prefix, a device can use stateless DHCPv6 to obtain other configuration parameters from a DHCPv6 server. This application is called stateless DHCPv6 configuration.

With an IPv6 address obtained through stateless address autoconfiguration, a device automatically enables the stateless DHCPv6 function after it receives an RA message with the managed address configuration flag (M flag) set to 0 and with the other stateful configuration flag (O flag) set to 1.

Stateless address autoconfiguration means that a node automatically generates an IPv6 address based on the information obtained through router/prefix discovery. For more information, see "Configuring IPv6 basics."

Operation

Figure 63 Operation of stateless DHCPv6



As shown in [Figure 63](#), stateless DHCPv6 operates in the following steps:

1. The DHCPv6 client multicasts an Information-request message to the multicast address of all DHCPv6 servers and DHCPv6 relay agents. The Information-request message contains an Option Request option, specifying the configuration parameters that the client requests from the DHCPv6 server.
2. After receiving the Information-request message, the DHCPv6 server returns the client a Reply message containing the requested configuration parameters.
3. The client checks the Reply message. If the obtained configuration parameters match those requested in the Information-request message, the client performs network configuration with the

parameters. If not, the client ignores the configuration parameters. If multiple replies are received, the first received reply will be used.

Protocols and standards

- RFC 3736, *Stateless Dynamic Host Configuration Protocol (DHCP) Service for IPv6*
- RFC 3315, *Dynamic Host Configuration Protocol for IPv6 (DHCPv6)*
- RFC 2462, *IPv6 Stateless Address Autoconfiguration*
- RFC 3633, *IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6*

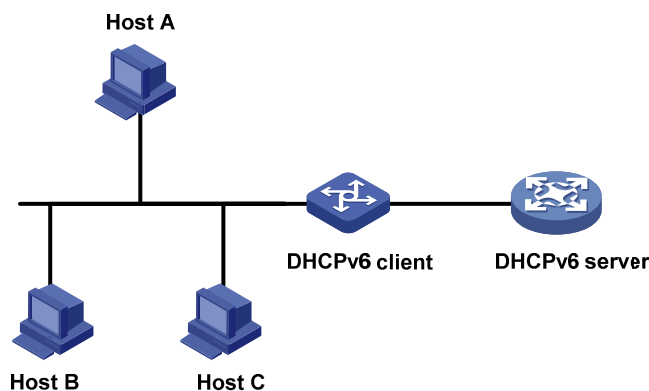
Configuring DHCPv6 server

Overview

As shown in [Figure 64](#), the DHCPv6 server assigns the DHCPv6 client an IPv6 prefix to facilitate IPv6 address management and network configuration. After obtaining the IPv6 prefix, the DHCPv6 client sends an RA message containing the prefix information to the subnet where it resides, so that hosts on the subnet can automatically configure their IPv6 addresses by using the prefix.

A device serving as a DHCPv6 server assigns DHCPv6 clients IPv6 prefixes, but not IPv6 addresses, and supports DHCPv6 stateless configuration to assign other configuration parameters.

Figure 64 Typical DHCPv6 server application



Concepts

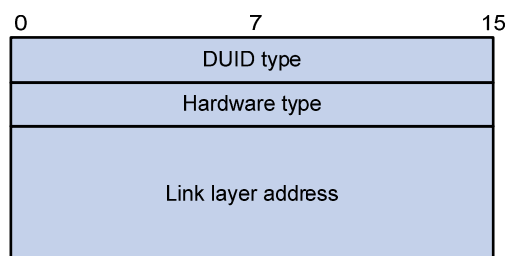
DHCPv6 multicast address

The multicast address FF05::1:3 identifies all DHCPv6 servers on the site-local network. The multicast address FF02::1:2 identifies all DHCPv6 servers and relay agents on the link-local link.

DUID

A DHCP unique identifier (DUID) uniquely identifies a DHCPv6 device (DHCPv6 client, server, or relay agent).

Figure 65 DUID-LL format



A DUID based on link-layer address (DUID-LL) defined in RFC 3315 is used to identify a DHCPv6 device. [Figure 65](#) shows the DUID-LL format, where:

- **DUID type**—The device supports DUID-LL as the DUID type with the value of 0x0003.
- **Hardware type**—The device supports Ethernet as the hardware type with the value of 0x0001.
- **Link layer address**—Its value is the bridge MAC address of the device.

IA

Identified by an IAID, an Identity Association (IA) provides a construct through which the obtained addresses, prefixes, and other configuration parameters assigned from a server to a client are managed. A client can maintain multiple IAs, each of which is configured on an interface to manage the addresses, prefixes, and other configuration parameters obtained by that interface.

IAID

An IAID uniquely identifies an IA. It is chosen by the client and must be unique among the IAIDs on the client.

PD

The Prefix Delegation (PD) is the lease record created by the DHCPv6 server for each assigned prefix. The PD contains information such as the IPv6 prefix, client DUID, IAID, valid lifetime, preferred lifetime, lease expiration time, and the IPv6 address of the requesting client.

Prefix selection process

Upon receiving a request, the DHCPv6 server selects the prefix and other configuration parameters from the address pool that is applied to the interface receiving the request. An address pool may contain the static prefixes configured for specific clients, or have a prefix pool referenced for dynamic assignment from the specific prefix range.

A DHCPv6 server selects a prefix from the address pool according to the following sequence:

1. The desired static prefix with the DUID and IAID matching those of the client
2. The static prefix with the DUID and IAID matching those of the client
3. The desired static prefix with the DUID matching the client's DUID and with no client IAID specified
4. The static prefix with the DUID matching the client's DUID and with no client IAID specified
5. The desired idle prefix in the prefix pool
6. An idle prefix in the prefix pool

DHCPv6 server configuration task list

Before you configure the DHCPv6 server, enable IPv6 by using the **ipv6** command.

Task	Remarks
Enabling the DHCPv6 server	Required
Creating a prefix pool	Required
Configuring a DHCPv6 address pool	Required
Applying the address pool to an interface	Required
Setting the DSCP value for DHCPv6 packets	Optional

Enabling the DHCPv6 server

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable the DHCPv6 server function.	ipv6 dhcp server enable	Disabled by default

Creating a prefix pool

A prefix pool specifies a range of prefixes.

To create a prefix pool:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a prefix pool.	ipv6 dhcp prefix-pool <i>prefix-pool-number</i> prefix <i>prefix/prefix-len</i> assign-len <i>assign-len</i>	Not configured by default

Configuring a DHCPv6 address pool

You can configure prefixes and other configuration parameters, such as the DNS server address, domain name, SIP server address, domain name of the SIP server, and address family translation router (AFTR) in a DHCPv6 address pool, for the DHCPv6 server to assign them to DHCPv6 clients.

Configuration restrictions and guidelines

- Only one prefix pool can be referenced by an address pool.
- A non-existing prefix pool can be referenced by an address pool. However, no prefix is available in the prefix pool for dynamic prefix assignment until the prefix pool is created.
- You cannot modify the prefix pool referenced by an address pool, or the preferred lifetime or valid lifetime by using the **prefix-pool** command. You must remove the configuration before you can have another prefix pool referenced by the address pool, or modify the preferred lifetime and valid lifetime.
- You can configure up to eight DNS server addresses, one domain name, eight SIP server addresses, and eight SIP server domain names in an address pool.

Configuration procedure

To configure a DHCPv6 address pool:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a DHCPv6 address pool and enter DHCPv6 address pool view.	ipv6 dhcp pool <i>pool-number</i>	Not configured by default.

Step	Command	Remarks
3. Configure a DHCPv6 address pool.	<ul style="list-style-type: none"> Configure a static prefix: static-bind prefix <i>prefix/prefix-len</i> duid <i>duid</i> [iaid <i>iaid</i>] [preferred-lifetime <i>preferred-lifetime</i> valid-lifetime <i>valid-lifetime</i>] Apply a prefix pool to the address pool: prefix-pool <i>prefix-pool-number</i> [preferred-lifetime <i>preferred-lifetime</i> valid-lifetime <i>valid-lifetime</i>] 	Use either command. No prefix is specified by default.
4. Configure a DNS server address.	dns-server <i>ipv6-address</i>	Optional. Not configured by default.
5. Configure a domain name.	domain-name <i>domain-name</i>	Optional. Not configured by default.
6. Configure the IPv6 address or domain name of a SIP server.	sip-server { address <i>ipv6-address</i> domain-name <i>domain-name</i> }	Optional. Not configured by default.
7. Specify the AFTR address.	ds-lite address <i>ipv6-address</i>	Optional. Not specified by default.

Applying the address pool to an interface

After an address pool is applied to an interface, a prefix and other configuration parameters can be selected from the address pool and assigned to the DHCPv6 client requesting through the interface.

Follow these guidelines when you apply an address pool to an interface:

- An interface cannot serve as a DHCPv6 server and DHCPv6 relay agent at the same time.
- It is not recommended that you enable DHCPv6 server and DHCPv6 client on the same interface.
- Only one address pool can be applied to an interface.
- A non-existing address pool can be applied to an interface. However, the server cannot assign any prefix or other configuration information from the address pool until the address pool is created.
- You cannot modify the address pool applied to an interface or parameters such as the server priority by using the **ipv6 dhcp server apply pool** command. You must remove the applied address pool before you can apply another address pool to the interface or modify parameters such as the server priority.

To apply an address pool to an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type interface-number</i>	N/A
3. Apply the DHCPv6 address pool to the interface.	ipv6 dhcp server apply pool <i>pool-number</i> [allow-hint preference <i>preference-value</i> rapid-commit] *	Not configured by default

Setting the DSCP value for DHCPv6 packets

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the DSCP value for DHCPv6 packets sent by the DHCPv6 server.	ipv6 dhcp dscp <i>dscp-value</i>	Optional. By default, the DSCP value in DHCPv6 packets is 56.

Displaying and maintaining the DHCPv6 server

Task	Command	Remarks
Display the DUID of the local device.	display ipv6 dhcp duid [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the DHCPv6 address pool information.	display ipv6 dhcp pool [<i>pool-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the prefix pool information.	display ipv6 dhcp prefix-pool [<i>prefix-pool-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the DHCPv6 server configuration information.	display ipv6 dhcp server [<i>interface interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the PD information.	display ipv6 dhcp server pd-in-use { all pool <i>pool-number</i> prefix <i>prefix/prefix-len</i> prefix-pool <i>prefix-pool-number</i> } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display packet statistics on the DHCPv6 server.	display ipv6 dhcp server statistics [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear PD information on the DHCPv6 server.	reset ipv6 dhcp server pd-in-use { all pool <i>pool-number</i> prefix <i>prefix/prefix-len</i> }	Available in user view
Clear packets statistics on the DHCPv6 server.	reset ipv6 dhcp server statistics	Available in user view

DHCPv6 server configuration example

Network requirements

As shown in [Figure 66](#), the switch serves as a DHCPv6 server, and assigns the IPv6 prefix, DNS server address, domain name, SIP server address, and SIP server domain name to the DHCPv6 clients. The IPv6 address of the switch is 1::1/64.

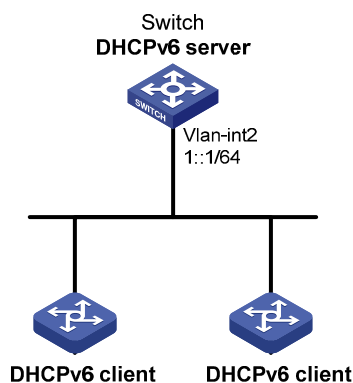
The switch assigns prefix 2001:0410:0201::/48 to the client whose DUID is 00030001CA0006A40000, and assigns prefixes ranging from 2001:0410::/48 to 2001:0410:FFFF::/48 (excluding 2001:0410:0201::/48) to other clients. The DNS server address is 2::2:3. The DHCPv6 clients reside in domain aaa.com. The SIP server address is 2:2::4, and the domain name of the SIP server is bbb.com.

Configuration considerations

To configure the DHCPv6 server:

- Enable IPv6 and DHCPv6 server.
- Create a prefix pool containing prefix 2001:0410::/32 with the length of the assigned prefix being 48, so that the server assigns clients the prefixes ranging 2001:0410::/48 to 2001:0410:FFFF::/48.
- Create an address pool. Configure a static prefix in the address pool and have the prefix pool referenced by the address pool. Configure other configuration parameters.
- Apply the address pool to the interface through which the server is connected to the clients.

Figure 66 Network diagram



Configuration procedure

Enable IPv6 and DHCPv6 server.

```
<Switch> system-view
[Switch] ipv6
[Switch] ipv6 dhcp server enable
```

Configure the IPv6 address of VLAN-interface 2.

```
[Switch] interface vlan-interface 2
[Switch-Vlan-interface2] ipv6 address 1::1/64
[Switch-Vlan-interface2] quit
```

Create and configure prefix pool 1.

```
[Switch] ipv6 dhcp prefix-pool 1 prefix 2001:0410::/32 assign-len 48
```

Create address pool 1.

```
[Switch] ipv6 dhcp pool 1
```

Apply prefix pool 1 to address pool 1, and set the preferred lifetime to one day, the valid lifetime to three days.

```
[Switch-ipv6-dhcp-pool-1] prefix-pool 1 preferred-lifetime 86400 valid-lifetime 259200
```

Configure static prefix 2001:0410:0201::/48 in address pool 1, and set the client DUID as 00030001CA0006A40000, the preferred lifetime to one day, and the valid lifetime to three days.

```
[Switch-ipv6-dhcp-pool-1] static-bind prefix 2001:0410:0201::/48 duid
00030001CA0006A40000 preferred-lifetime 86400 valid-lifetime 259200
```

Configure the DNS server address as 2:2::3.

```

[Switch-ipv6-dhcp-pool-1] dns-server 2:2::3
# Configure the domain name as aaa.com.
[Switch-ipv6-dhcp-pool-1] domain-name aaa.com
# Configure the SIP server address as 2:2::4, and the domain name of the SIP server as bbb.com.
[Switch-ipv6-dhcp-pool-1] sip-server address 2:2::4
[Switch-ipv6-dhcp-pool-1] sip-server domain-name bbb.com
[Switch-ipv6-dhcp-pool-1] quit
# Apply address pool 1 to VLAN-interface 2, configure the address pool to support the desired prefix
assignment and rapid prefix assignment, and set the precedence to the highest.
[Switch] interface vlan-interface 2
[Switch-Vlan-interface2] ipv6 dhcp server apply pool 1 allow-hint preference 255
rapid-commit

```

Verifying the configuration

```

# Display the DHCPv6 server configuration information on VLAN-interface 2.
[Switch-Vlan-interface2] display ipv6 dhcp server interface vlan-interface 2
Using pool: 1
Preference value: 255
Allow-hint: Enabled
Rapid-commit: Enabled
# Display the information of address pool 1.
[Switch-Vlan-interface2] display ipv6 dhcp pool 1
DHCPv6 pool: 1
  Static bindings:
    DUID: 00030001CA0006A40000
    IAID: A1A1A1A1
    Prefix: 2001:410:201::/48
      preferred lifetime 86400, valid lifetime 2592000
  Prefix pool: 1
    preferred lifetime 86400, valid lifetime 2592000
  DNS server address:
    2:2::3
  Domain name: aaa.com
  SIP server address:
    2:2::4
  SIP server domain name:
    bbb.com
# Display the information of prefix pool 1.
[Switch-Vlan-interface2] display ipv6 dhcp prefix-pool 1
Prefix: 2001:410::/32
Assigned length: 48
Total prefix number: 65536
Available: 65535
In-use: 0
Static: 1

```

After the client whose DUID is 00030001CA0006A40000 obtains an IPv6 prefix, display the PD information on the DHCPv6 server.

```
[Switch-Vlan-interface2] display ipv6 dhcp server pd-in-use all
```

Total number = 1

Prefix	Type	Pool	Lease-expiration
2001:410:201::/48	Static(C)	1	Jul 10 2009 19:45:01

After the other client obtains an IPv6 prefix, display the PD information on the DHCPv6 server.

```
[Switch-Vlan-interface2] display ipv6 dhcp server pd-in-use all
```

Total number = 2

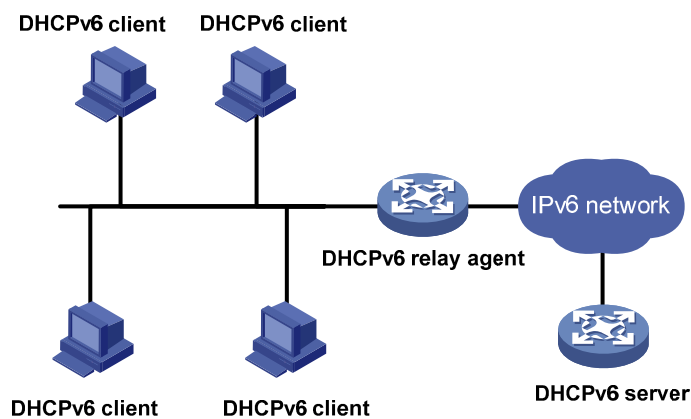
Prefix	Type	Pool	Lease-expiration
2001:410:201::/48	Static(C)	1	Jul 10 2009 19:45:01
2001:410::/48	Auto(C)	1	Jul 10 2009 20:44:05

Configuring DHCPv6 relay agent

Overview

A DHCPv6 client usually uses a multicast address to contact the DHCPv6 server on the local link to obtain an IPv6 address and other configuration parameters. As shown in [Figure 67](#), if the DHCPv6 server resides on another subnet, the DHCPv6 client can contact the server via a DHCPv6 relay agent, so you do not need to deploy a DHCPv6 server on each subnet.

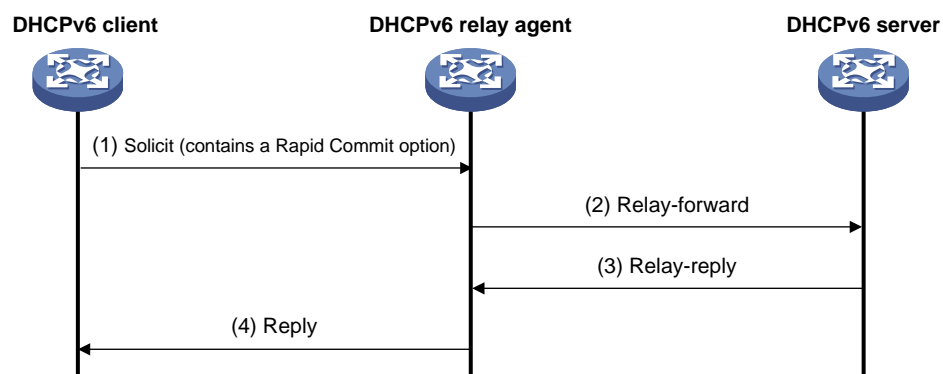
Figure 67 Typical DHCPv6 relay agent application



DHCPv6 relay agent operation

[Figure 68](#) shows how the DHCPv6 client obtains an IPv6 address and other network configuration parameters from the DHCPv6 server through the DHCPv6 relay agent, using the process of rapid assignment involving two messages.

Figure 68 Operating process of a DHCPv6 relay agent



The operation process is as follows:

1. The DHCPv6 client sends a Solicit message containing the Rapid Commit option to the multicast address FF02::1:2 of all the DHCPv6 servers and relay agents.

2. After receiving the Solicit message, the DHCPv6 relay agent encapsulates the message into the Relay Message option of a Relay-forward message, and sends the message to the DHCPv6 server.
3. After obtaining the Solicit message from the Relay-forward message, the DHCPv6 server selects an IPv6 address and other required parameters, and adds them to the reply which is encapsulated within the Relay Message option of a Relay-reply message. The DHCPv6 server then sends the Relay-reply message to the DHCPv6 relay agent.
4. The DHCPv6 relay agent obtains the reply from the Relay-reply message and sends the reply to the DHCPv6 client.

The DHCPv6 client uses the IPv6 address and other network parameters assigned by the DHCPv6 server to perform network configuration.

Configuring the DHCPv6 relay agent

Upon receiving a Solicit message from a DHCPv6 client, the interface that operates as a DHCPv6 relay agent encapsulates the request into a Relay-forward message and forwards the message to the specified DHCPv6 server, which then assigns an IPv6 address and other configuration parameters to the DHCPv6 client.

Configuration guidelines

Follow these guidelines when you configure the DHCPv6 relay agent:

- Before you configure the DHCPv6 relay agent, enable IPv6 by using the **ipv6** command in system view.
- Executing the **ipv6 dhcp relay server-address** command repeatedly can specify multiple DHCPv6 servers. Up to eight DHCPv6 servers can be specified for an interface. After receiving requests from DHCPv6 clients, the DHCPv6 relay agent forwards the requests to all the specified DHCPv6 servers.
- If the DHCPv6 server address is a link-local address or link-scoped multicast address on the local link, you must specify an outgoing interface using the **interface** keyword in the **ipv6 dhcp relay server-address** command. Otherwise, DHCPv6 packets may fail to be forwarded to the DHCPv6 server.
- After you remove all the specified DHCPv6 servers from an interface with the **undo ipv6 dhcp relay server-address** command, DHCPv6 relay agent is disabled on the interface.
- An interface cannot serve as a DHCPv6 relay agent and DHCPv6 server at the same time.
- HP does not recommend enabling the DHCPv6 relay agent and DHCPv6 client on the same interface

Configuration procedure

To configure the DHCPv6 relay agent:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A

Step	Command	Remarks
3. Enable DHCPv6 relay agent on the interface and specify a DHCPv6 server.	ipv6 dhcp relay server-address <i>ipv6-address [interface interface-type interface-number]</i>	By default, DHCPv6 relay agent is disabled and no DHCPv6 server is specified on the interface.

Setting the DSCP value for DHCPv6 packets

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the DSCP value for DHCPv6 packets sent by the DHCPv6 relay agent.	ipv6 dhcp dscp <i>dscp-value</i>	Optional. By default, the DSCP value in DHCPv6 packets is 56.

Displaying and maintaining the DHCPv6 relay agent

Task	Command	Remarks
Display the DUID of the local device.	display ipv6 dhcp duid [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display DHCPv6 server addresses specified on the DHCPv6 relay agent.	display ipv6 dhcp relay server-address { all interface <i>interface-type interface-number</i> } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display packet statistics on the DHCPv6 relay agent.	display ipv6 dhcp relay statistics [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear packets statistics on the DHCPv6 relay agent.	reset ipv6 dhcp relay statistics	Available in user view

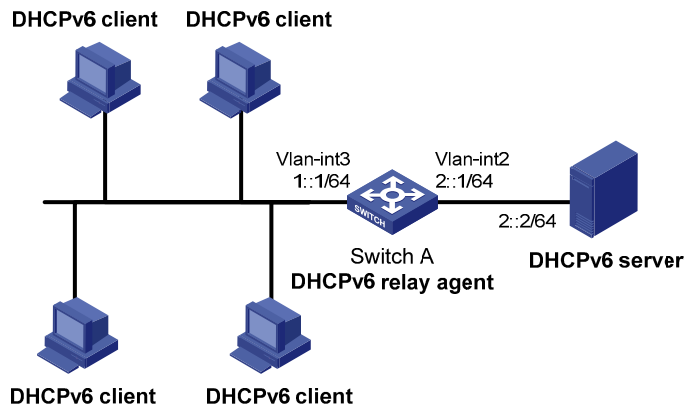
DHCPv6 relay agent configuration example

Network requirements

As shown in [Figure 69](#), the network address prefix of DHCPv6 clients is 1::/64, and the IPv6 address of the DHCPv6 server is 2::2/64. The DHCPv6 client and server need to communicate via a DHCPv6 relay agent (Switch A).

Switch A acts as the gateway of network 1::/64. It sends RA messages to notify the hosts to obtain IPv6 addresses and other configuration parameters through DHCPv6.

Figure 69 Network diagram



Configuration procedure

1. Configure Switch A as a DHCPv6 relay agent:

Enable the IPv6 packet forwarding function.

```
<SwitchA> system-view
[SwitchA] ipv6
```

Configure the IPv6 addresses of VLAN-interface 2 and VLAN-interface 3 respectively.

```
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ipv6 address 2::1 64
[SwitchA-Vlan-interface2] quit
[SwitchA] interface vlan-interface 3
[SwitchA-Vlan-interface3] ipv6 address 1::1 64
```

Enable DHCPv6 relay agent and specify the DHCPv6 server address on VLAN-interface 3.

```
[SwitchA-Vlan-interface3] ipv6 dhcp relay server-address 2::2
```

2. Configure Switch A as a gateway:

Enable Switch A to send RA messages and turn on the M and O flags.

```
[SwitchA-Vlan-interface3] undo ipv6 nd ra halt
[SwitchA-Vlan-interface3] ipv6 nd autoconfig managed-address-flag
[SwitchA-Vlan-interface3] ipv6 nd autoconfig other-flag
```

Verifying the configuration

Display DHCPv6 server address information on Switch A.

```
[SwitchA-Vlan-interface3] display ipv6 dhcp relay server-address all
Interface: Vlan3
Server address(es)                               Output Interface
2::2
```

Display packet statistics on the DHCPv6 relay agent.

```
[SwitchA-Vlan-interface3] display ipv6 dhcp relay statistics
Packets dropped           : 0
Error                     : 0
Excess of rate limit     : 0
Packets received         : 14
```


SOLICIT	:	0
REQUEST	:	0
CONFIRM	:	0
RENEW	:	0
REBIND	:	0
RELEASE	:	0
DECLINE	:	0
INFORMATION-REQUEST	:	7
RELAY-FORWARD	:	0
RELAY-REPLY	:	7
Packets sent	:	14
ADVERTISE	:	0
RECONFIGURE	:	0
REPLY	:	7
RELAY-FORWARD	:	7
RELAY-REPLY	:	0

Configuring DHCPv6 client

Overview

Serving as a DHCPv6 client, the device only supports stateless DHCPv6 configuration, that is, the device can only obtain other network configuration parameters, except the IPv6 address and prefix from the DHCPv6 server.

With an IPv6 address obtained through stateless address autoconfiguration, the device automatically enables the stateless DHCPv6 function after it receives an RA message with the M flag set to 0 and the O flag set to 1.

Configuring the DHCPv6 client

Configuration prerequisites

To make the DHCPv6 client successfully obtain configuration parameters through stateless DHCPv6 configuration, make sure that the DHCPv6 server is available.

Configuration guidelines

- For more information about the **ipv6 address auto** command, see the *Layer 3—IP Services Command Reference*.
- HP does not recommend enabling the DHCPv6 client and DHCPv6 server, or the DHCPv6 client and DHCPv6 relay agent on the same interface at the same time.

Configuration procedure

To configure the DHCPv6 client:

Step	Command
1. Enter system view.	system-view
2. Enable the IPv6 packet forwarding function.	ipv6
3. Enter interface view.	interface <i>interface-type interface-number</i>
4. Enable IPv6 stateless address autoconfiguration.	ipv6 address auto

Setting the DSCP value for DHCPv6 packets

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Set the DSCP value for the DHCPv6 packets sent by the DHCPv6 client.	<code>ipv6 dhcp client dscp dscp-value</code>	Optional. By default, the DSCP value in DHCPv6 packets is 56.

Displaying and maintaining the DHCPv6 client

Task	Command	Remarks
Display DHCPv6 client information.	<code>display ipv6 dhcp client [interface interface-type interface-number] [{ begin exclude include } regular-expression]</code>	Available in any view
Display DHCPv6 client statistics.	<code>display ipv6 dhcp client statistics [interface interface-type interface-number] [{ begin exclude include } regular-expression]</code>	Available in any view
Display the DUID of the local device.	<code>display ipv6 dhcp duid [{ begin exclude include } regular-expression]</code>	Available in any view
Clear DHCPv6 client statistics.	<code>reset ipv6 dhcp client statistics [interface interface-type interface-number]</code>	Available in user view

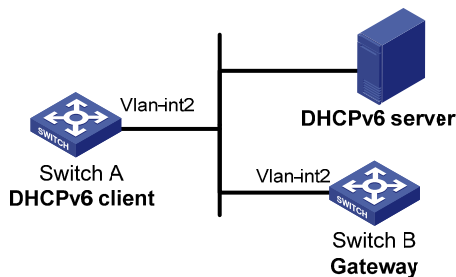
Stateless DHCPv6 configuration example

Network requirements

As shown in [Figure 70](#), through stateless DHCPv6, Switch A obtains the DNS server address, domain name, and other information from the server.

Switch B acts as the gateway to send RA messages periodically.

Figure 70 Network diagram



Configuration procedure

- Configure Switch B:
 - # Enable the IPv6 packet forwarding function.
 - <SwitchB> system-view
 - [SwitchB] ipv6
 - # Configure the IPv6 address of VLAN-interface 2.

```
[SwitchB] interface vlan-interface 2
[SwitchB-Vlan-interface2] ipv6 address 1::1 64
# Set the O flag in the RA messages to 1.
[SwitchB-Vlan-interface2] ipv6 nd autoconfig other-flag
# Enable Switch B to send RA messages.
[SwitchB-Vlan-interface2] undo ipv6 nd ra halt
```

2. Configure Switch A:

Enable the IPv6 packet forwarding function.

```
<SwitchA> system-view
[SwitchA] ipv6
```

Enable stateless IPv6 address autoconfiguration on VLAN-interface 2.

```
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ipv6 address auto
```

With this command executed, if VLAN-interface 2 has no IPv6 address configured, Switch A will automatically generate a link-local address, and send an RS message, requesting the gateway (Switch B) to reply with an RA message immediately.

Verifying the configuration

After receiving an RA message with the M flag set to 0 and the O flag set to 1, Switch A automatically enables the stateless DHCPv6 function.

Use the **display ipv6 dhcp client** command to view the current client configuration information. If the client successfully obtains configuration information from the server, the following information will be displayed.

```
[SwitchA-Vlan-interface2] display ipv6 dhcp client interface vlan-interface 2
Vlan-interface2 is in stateless DHCPv6 client mode
State is OPEN
Preferred Server:
  Reachable via address      : FE80::213:7FFF:FEF6:C818
  DUID                       : 0003000100137ff6c818
  DNS servers                : 1:2:3:5
                             1:2:4:7
  Domain names               : abc.com
                             Sysname.com
```

Use the **display ipv6 dhcp client statistics** command to view the current client statistics.

```
[SwitchA-Vlan-interface2] display ipv6 dhcp client statistics
Interface                : Vlan-interface2
Packets Received         : 1
  Reply                   : 1
  Advertise                : 0
  Reconfigure              : 0
  Invalid                  : 0
Packets Sent             : 5
  Solicit                  : 0
  Request                  : 0
  Confirm                  : 0
  Renew                    : 0
```

Rebind	:	0
Information-request	:	5
Release	:	0
Decline	:	0

Configuring DHCPv6 snooping

A DHCPv6 snooping device does not work if it is between a DHCPv6 relay agent and a DHCPv6 server. The DHCPv6 snooping device works when it is between a DHCPv6 client and a DHCPv6 relay agent or between a DHCPv6 client and a DHCPv6 server.

You can configure only Layer 2 Ethernet ports or Layer 2 aggregate interfaces as DHCPv6 snooping trusted ports. For more information about aggregate interfaces, see *Layer 2—LAN Switching Configuration Guide*.

Overview

DHCPv6 snooping is security feature with the following functions:

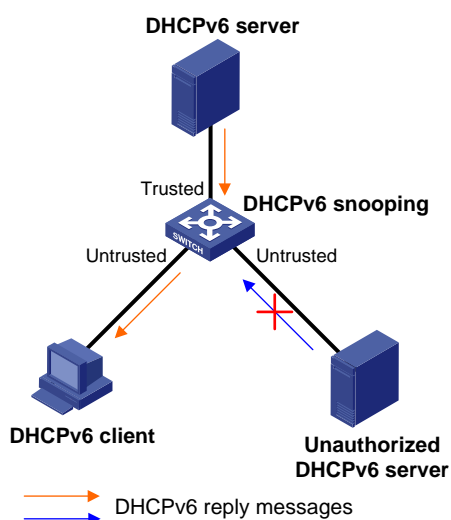
- Ensure that DHCPv6 clients obtain IPv6 addresses from authorized DHCPv6 servers.
- Record IP-to-MAC mappings of DHCPv6 clients.

Ensuring that DHCPv6 clients obtain IPv6 addresses from authorized DHCPv6 servers

If DHCPv6 clients obtain invalid IPv6 addresses and network configuration parameters from an unauthorized DHCP server, they will be unable to communicate normally with other network devices. With DHCPv6 snooping, the ports of a device can be configured as trusted or untrusted to make sure that the clients obtain IPv6 addresses only from authorized DHCPv6 servers.

- **Trusted**—A trusted port forwards DHCPv6 messages normally.
- **Untrusted**—An untrusted port discards reply messages from any DHCPv6 server.

Figure 71 Trusted and untrusted ports



A DHCPv6 snooping device's port that is connected to an authorized DHCPv6 server, DHCPv6 relay agent, or another DHCPv6 snooping device should be configured as a trusted port. The trusted port forwards reply messages from the authorized DHCPv6 server. Other ports are configured as untrusted so

that they do not forward reply messages from any DHCPv6 servers. This ensures that the DHCPv6 client can obtain an IPv6 address from the authorized DHCPv6 server only.

As shown in [Figure 71](#), configure the port that connects to the DHCPv6 server as a trusted port, and other ports as untrusted.

Recording IP-to-MAC mappings of DHCPv6 clients

DHCPv6 snooping reads DHCPv6 messages to create and update DHCPv6 snooping entries, including MAC addresses of clients, IPv6 addresses obtained by the clients, ports that connect to DHCPv6 clients, and VLANs to which the ports belong. You can use the **display ipv6 dhcp snooping user-binding** command to view the IPv6 address obtained by each client, so you can manage and monitor the clients' IPv6 addresses.

Enabling DHCPv6 snooping

To allow clients to obtain IPv6 addresses from an authorized DHCPv6 server, enable DHCPv6 snooping globally and configure trusted and untrusted ports properly. To record DHCPv6 snooping entries for a VLAN, enable DHCPv6 snooping for the VLAN.

To enable DHCPv6 snooping:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable DHCPv6 snooping globally.	ipv6 dhcp snooping enable	Disabled by default.
3. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
4. Enable DHCPv6 snooping for the VLAN.	ipv6 dhcp snooping vlan enable	Optional. Disabled by default.

Configuring a DHCPv6 snooping trusted port

After enabling DHCPv6 snooping globally, you can specify trusted and untrusted ports for a VLAN as needed. A DHCPv6 snooping trusted port normally forwards received DHCPv6 packets. A DHCPv6 snooping untrusted port discards any DHCPv6 reply message received from a DHCPv6 server. Upon receiving a DHCPv6 request from a client in the VLAN, the DHCPv6 snooping device forwards the packet through trusted ports rather than any untrusted port in the VLAN, reducing network traffic.

You must specify a port connected to an authorized DHCPv6 server as trusted to make sure that DHCPv6 clients can obtain valid IPv6 addresses. The trusted port and the ports connected to the DHCPv6 clients must be in the same VLAN.

If a Layer 2 Ethernet port is added to an aggregation group, the DHCPv6 snooping configuration of the interface will not take effect until the interface quits from the aggregation group.

To configure a DHCPv6 snooping trusted port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the port as trusted.	ipv6 dhcp snooping trust	By default, all ports of the device with DHCPv6 snooping globally enabled are untrusted.

Configuring the maximum number of DHCPv6 snooping entries an interface can learn

Perform this optional task to prevent an interface from learning too many DHCPv6 snooping entries and to save system resources.

To configure the maximum number of DHCPv6 snooping entries an interface can learn:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the maximum number of DHCPv6 snooping entries that the interface can learn.	ipv6 dhcp snooping max-learning-num <i>number</i>	Optional. By default, the number of DHCPv6 snooping entries learned by an interface is not limited.

Displaying and maintaining DHCPv6 snooping

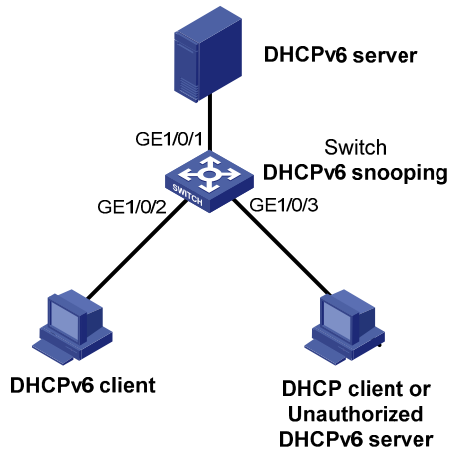
Task	Command	Remarks
Display DHCPv6 snooping trusted ports.	display ipv6 dhcp snooping trust [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display DHCPv6 snooping entries.	display ipv6 dhcp snooping user-binding { <i>ipv6-address</i> dynamic } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear DHCPv6 snooping entries.	reset ipv6 dhcp snooping user-binding { <i>ipv6-address</i> dynamic }	Available in user view

DHCPv6 snooping configuration example

Network requirements

As shown in [Figure 72](#), Switch is connected to a DHCPv6 server through GigabitEthernet 1/0/1, and is connected to DHCPv6 clients through GigabitEthernet 1/0/2 and GigabitEthernet 1/0/3. These three interfaces belong to VLAN 2. Configure Switch to forward DHCPv6 reply messages received on GigabitEthernet 1/0/1 only and record the IP-to-MAC mappings for DHCPv6 clients.

Figure 72 Network diagram



Configuration procedure

Enable DHCPv6 snooping globally.

```
<Switch> system-view
```

```
[Switch] ipv6 dhcp snooping enable
```

Add GigabitEthernet 1/0/1, GigabitEthernet 1/0/2, and GigabitEthernet 1/0/3 to VLAN 2.

```
[Switch] vlan 2
```

```
[Switch-vlan2] port GigabitEthernet 1/0/1 GigabitEthernet 1/0/2 GigabitEthernet 1/0/3
```

Enable DHCPv6 snooping for VLAN 2.

```
[Switch-vlan2] ipv6 dhcp snooping vlan enable
```

```
[Switch] quit
```

Configure GigabitEthernet 1/0/1 as a DHCPv6 snooping trusted port.

```
[Switch] interface GigabitEthernet 1/0/1
```

```
[Switch-GigabitEthernet1/0/1] ipv6 dhcp snooping trust
```

Verifying the configuration

Connect GigabitEthernet 1/0/2 to a DHCPv6 client, GigabitEthernet 1/0/1 to a DHCPv6 server, and GigabitEthernet 1/0/3 to an unauthorized DHCPv6 server. The DHCPv6 client obtains an IPv6 address from DHCPv6 server, but cannot obtain any IPv6 address from the unauthorized DHCPv6 server. You can use the **display ipv6 dhcp snooping user-binding** command to view the DHCPv6 snooping entries on Switch.

Configuring IPv6 DNS

Overview

IPv6 Domain Name System (DNS) is responsible for translating domain names into IPv6 addresses. Like IPv4 DNS, IPv6 DNS includes static domain name resolution and dynamic domain name resolution. The functions and implementations of the two types of domain name resolution are the same as those of IPv4 DNS. For more information, see "Configuring IPv4 DNS."

Configuring the IPv6 DNS client

Configuring static domain name resolution

Configuring static domain name resolution refers to specifying the mappings between host names and IPv6 addresses. Static domain name resolution allows applications such as Telnet to contact hosts by using host names instead of IPv6 addresses.

Follow these guidelines when you configure static domain name resolution:

- A host name can be mapped to one IPv6 address only. If you map a host name to different IPv6 addresses, the last configuration takes effect.
- You can configure up to 50 mappings between domain name and IPv6 address on the switch.

To configure static domain name resolution:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure a mapping between a host name and an IPv6 address.	ipv6 host <i>hostname ipv6-address</i>	Not configured by default

Configuring dynamic domain name resolution

To send DNS queries to a correct server for resolution, dynamic domain name resolution needs to be enabled and a DNS server needs to be configured.

In addition, you can configure a DNS suffix that the system automatically adds to the provided domain name for resolution.

Follow these guidelines when you configure dynamic domain name resolution:

- You can configure up to six DNS servers, including those with IPv4 addresses on the switch.
- You can specify up to ten DNS suffixes on the switch.

To configure dynamic domain name resolution:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable dynamic domain name resolution.	dns resolve	Disabled by default.
3. Specify a DNS server.	dns server ipv6 <i>ipv6-address</i> [<i>interface-type interface-number</i>]	Not specified by default. If the IPv6 address of a DNS server is a link-local address, you must specify the <i>interface-type</i> and <i>interface-number</i> arguments.
4. Configure a DNS suffix.	dns domain <i>domain-name</i>	Optional. Not configured by default. Only the provided domain name is resolved.

Setting the DSCP value for IPv6 DNS packets

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the DSCP value for IPv6 DNS packets.	dns ipv6 dscp <i>dscp-value</i>	Optional. By default, the DSCP value in IPv6 DNS packets is 0.

Displaying and maintaining IPv6 DNS

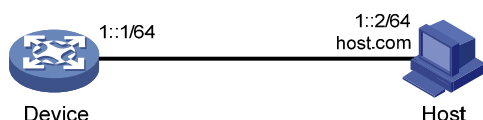
Task	Command	Remarks
Display the static IPv6 domain name resolution table.	display ipv6 host [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 DNS server information.	display dns ipv6 server [dynamic] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display DNS suffixes.	display dns domain [dynamic] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about dynamic IPv6 domain name cache.	display dns host ipv6 [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear information about dynamic IPv6 domain name cache.	reset dns host ipv6	Available in user view

Static domain name resolution configuration example

Network requirements

As shown in [Figure 73](#), the device wants to access the host by using an easy-to-remember domain name rather than an IPv6 address. Configure static domain name resolution on the device so that the device can use the domain name `host.com` to access the host whose IPv6 address is `1::2`.

Figure 73 Network diagram



Configuration procedure

Configure a mapping between host name `host.com` and IPv6 address `1::2`.

```
<Device> system-view
[Device] ipv6 host host.com 1::2
```

Enable IPv6 packet forwarding.

```
[Device] ipv6
```

Use the **ping ipv6 host.com** command to verify that the device can use static domain name resolution to resolve domain name `host.com` into IPv6 address `1::2`.

```
[Device] ping ipv6 host.com
PING host.com (1::2):
 56 data bytes, press CTRL_C to break
  Reply from 1::2
  bytes=56 Sequence=1 hop limit=128 time = 3 ms
  Reply from 1::2
  bytes=56 Sequence=2 hop limit=128 time = 1 ms
  Reply from 1::2
  bytes=56 Sequence=3 hop limit=128 time = 1 ms
  Reply from 1::2
  bytes=56 Sequence=4 hop limit=128 time = 2 ms
  Reply from 1::2
  bytes=56 Sequence=5 hop limit=128 time = 2 ms
--- host.com ping statistics ---
 5 packet(s) transmitted
 5 packet(s) received
 0.00% packet loss
 round-trip min/avg/max = 1/1/3 ms
```

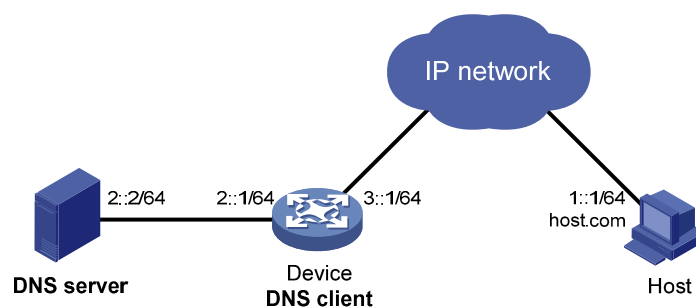
Dynamic domain name resolution configuration example

Network requirements

As shown in [Figure 74](#), the device wants to access the host by using an easy-to-remember domain name rather than an IPv6 address. The IPv6 address of the DNS server is `2::2/64` and the server has a `com` domain, which stores the mapping between domain name `host` and IPv6 address `1::1/64`.

Configure dynamic domain name resolution and the domain name suffix `com` on the device that serves as a DNS client so that the device can use domain name `host` to access the host with the domain name `host.com` and the IPv6 address `1::1/64`.

Figure 74 Network diagram



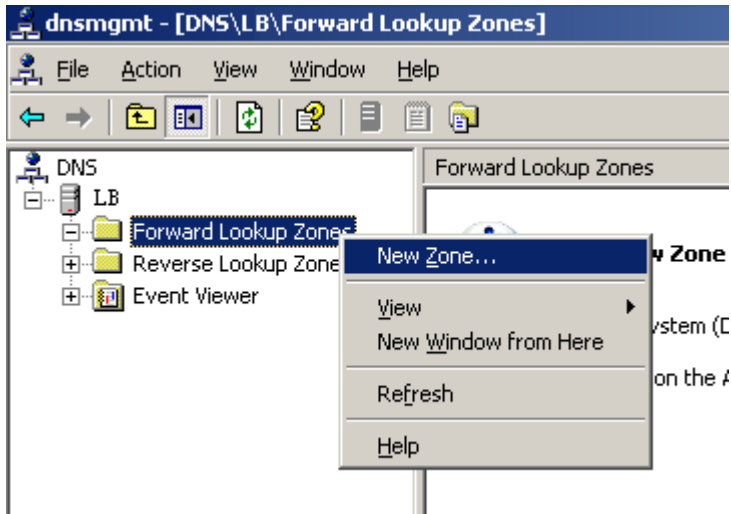
Configuration procedure

Before performing the following configuration, make sure that the device and the host are accessible to each other via available routes, and the IPv6 addresses of the interfaces are configured as shown [Figure 74](#).

This configuration may vary with DNS servers. The following configuration is performed on a PC running Windows Server 2003. Make sure that the DNS server supports the IPv6 DNS function so that the server can process IPv6 DNS packets, and the interfaces of the DNS server can forward IPv6 packets.

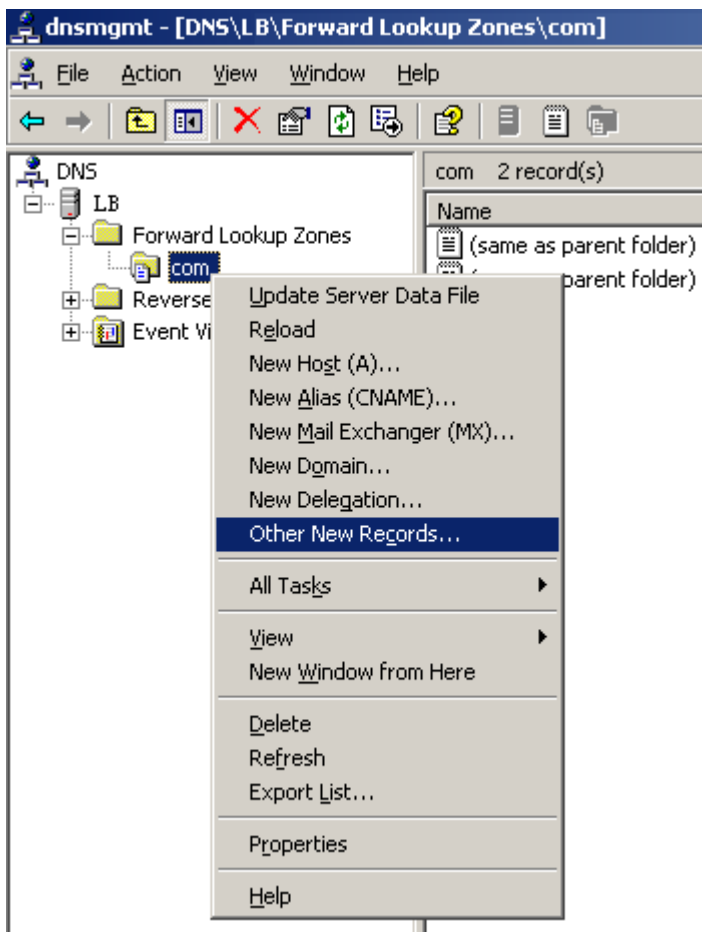
1. Configure the DNS server:
 - a. Select **Start > Programs > Administrative Tools > DNS**.
The DNS server configuration page appears, as shown in [Figure 75](#).
 - b. Right-click **Forward Lookup Zones**, select **New Zone**, and then follow the instructions to create a new zone named **com**.

Figure 75 Creating a zone



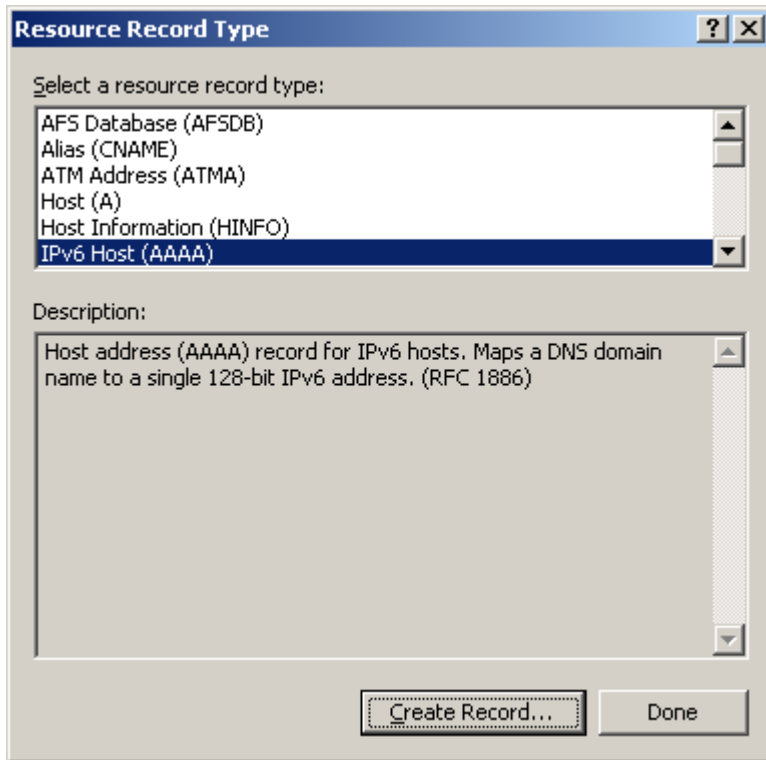
- c. On the DNS server configuration page, right-click zone **com** and select **Other New Records**.

Figure 76 Creating a record



- d. On the page that appears, select **IPv6 Host (AAAA)** as the resource record type, and click **Create Record**.

Figure 77 Selecting the resource record type



- e. On the page that appears, enter host name **host** and IPv6 address **1::1**.
 - f. Click **OK**.
- The mapping between the IP address and host name is created.

Figure 78 Adding a mapping between domain name and IPv6 address

The screenshot shows a window titled "New Resource Record" with a tab labeled "IPv6 Host (AAAA)". Inside the window, there are three text input fields. The first is labeled "Host (uses parent domain if left blank):" and contains the text "host". The second is labeled "Fully qualified domain name (FQDN):" and contains "host.com.". The third is labeled "IP version 6 host address:" and contains "1::1". At the bottom of the window, there are two buttons: "OK" and "Cancel".

2. Configure the DNS client:

Enable dynamic domain name resolution.

```
<Device> system-view
```

```
[Device] dns resolve
```

Specify the DNS server 2::2.

```
[Device] dns server ipv6 2::2
```

Configure com as the DNS suffix.

```
[Device] dns domain com
```

Verifying the configuration

Use the **ping ipv6 host** command on the device to verify that the communication between the device and the host is normal and that the corresponding destination IP address is 1::1.

```
[Device] ping ipv6 host
```

```
Trying DNS resolve, press CTRL_C to break
```

```
Trying DNS server (2::2)
```

```
PING host.com (1::1):
```

```
56 data bytes, press CTRL_C to break
```

```
Reply from 1::1
```

```
bytes=56 Sequence=1 hop limit=126 time = 2 ms
```

```
Reply from 1::1
```



```
bytes=56 Sequence=2 hop limit=126 time = 1 ms
Reply from 1::1
bytes=56 Sequence=3 hop limit=126 time = 1 ms
Reply from 1::1
bytes=56 Sequence=4 hop limit=126 time = 1 ms
Reply from 1::1
bytes=56 Sequence=5 hop limit=126 time = 1 ms
```

```
--- host.com ping statistics ---
```

```
5 packet(s) transmitted
```

```
5 packet(s) received
```

```
0.00% packet loss
```

```
round-trip min/avg/max = 1/1/2 ms
```

Configuring tunneling(only available on the HP 5500 EI)

Overview

Tunneling is an encapsulation technology: one network protocol encapsulates packets of another network protocol and transfers them over a virtual point-to-point connection. The virtual connection is called a tunnel. Packets are encapsulated and de-encapsulated at both ends of a tunnel. Tunneling refers to the whole process from data encapsulation to data transfer to data de-encapsulation.

Tunneling provides the following features:

- Transition techniques, such as IPv6 over IPv4 tunneling, to interconnect IPv4 and IPv6 networks.
- Virtual Private Networks (VPNs) for guaranteeing communication security, such as IPv4 over IPv4 tunneling, IPv4/IPv6 over IPv6 tunneling, Generic Routing Encapsulation (GRE), Dynamic Virtual Private Network (DVPN), and IPsec tunneling.
- Traffic engineering, such as Multiprotocol Label Switching traffic engineering (MPLS TE) to prevent network congestion.

Unless otherwise specified, the term "tunnel" used throughout this chapter refers to an IPv6 over IPv4, IPv4 over IPv4, IPv4 over IPv6, or IPv6 over IPv6 tunnel.

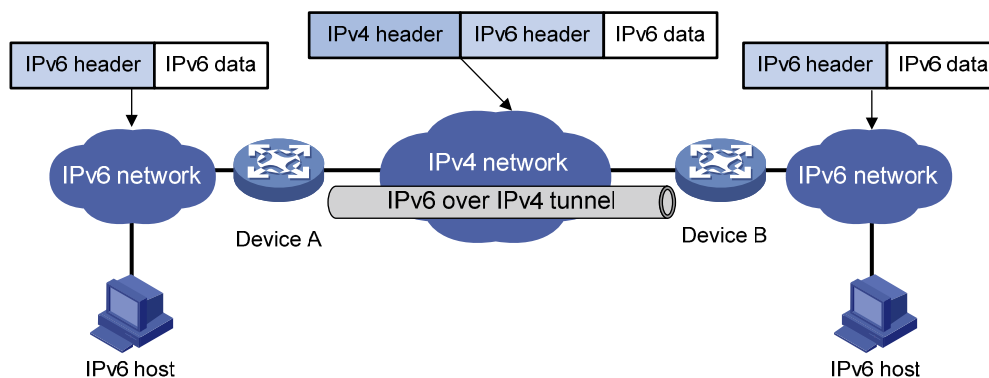
IPv6 over IPv4 tunneling

Implementation

IPv6 over IPv4 tunneling adds an IPv4 header to IPv6 data packets so that IPv6 packets can pass an IPv4 network through a tunnel to realize internetworking between isolated IPv6 networks, as shown in [Figure 79](#). The IPv6 over IPv4 tunnel can be established between two hosts, a host and a device, or two devices. The tunnel destination node can forward IPv6 packets if it is not the destination of the IPv6 packets.

The devices at both ends of an IPv6 over IPv4 tunnel must support the IPv4/IPv6 dual stack.

Figure 79 IPv6 over IPv4 tunnel



The IPv6 over IPv4 tunnel processes packets in the following ways.

1. A host in the IPv6 network sends an IPv6 packet to Device A at the tunnel source.

2. After determining from the routing table that the packet needs to be forwarded through the tunnel, Device A encapsulates the IPv6 packet with an IPv4 header and forwards it through the physical interface of the tunnel.
3. Upon receiving the packet, Device B de-encapsulates the packet.
4. Device B forwards the packet according to the destination address in the de-encapsulated IPv6 packet. If the destination address is the device itself, Device B forwards the IPv6 packet to the upper-layer protocol for processing.

Tunnel types

Depending on how the IPv4 address of the tunnel destination is acquired, IPv6 over IPv4 tunnels are divided into the following types:

- **Manually configured tunnel**—The destination address of the tunnel cannot be automatically acquired through the destination IPv6 address of an IPv6 packet at the tunnel source, and must be manually configured.
- **Automatic tunnel**—The destination address of the tunnel is an IPv6 address with an IPv4 address embedded, and the IPv4 address can be automatically acquired through the destination IPv6 address of an IPv6 packet at the tunnel source.

Table 10 IPv6 over IPv4 tunnel modes and key parameters

Tunnel type	Tunnel mode	Tunnel source/destination address	Tunnel interface address type
Manually configured tunnel	IPv6 manual tunneling	The source/destination IP address is a manually configured IPv4 address.	IPv6 address
	6to4 tunneling	The source IP address is a manually configured IPv4 address. The destination IP address does not need to be configured.	6to4 address, in the format of 2002:IPv4-source-address::/48
Automatic tunnel	Intra-site automatic tunnel addressing protocol (ISATAP) tunneling	The source IP address is a manually configured IPv4 address. The destination IP address does not need to be configured.	ISATAP address, in the format of Prefix:0:5EFE:IPv4-source-address/64

According to the way an IPv6 packet is encapsulated, IPv6 over IPv4 tunnels are divided into the following modes:

- IPv6 manual tunneling

A manually configured tunnel is a point-to-point link. Each link is a separate tunnel. IPv6 manual tunnels are mainly used to provide stable connections for regular secure communication between border routers or between border routers and hosts for access to remote IPv6 networks.
- 6to4 tunneling

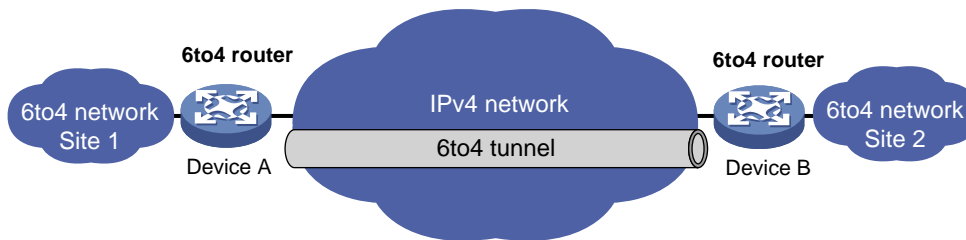
An automatic 6to4 tunnel is a point-to-multipoint tunnel mainly constructed between edge routers, and is used to connect multiple isolated IPv6 networks over an IPv4 network to remote IPv6 networks. The embedded IPv4 address in an IPv6 address is used to automatically acquire the destination IPv4 address of the tunnel.

The automatic 6to4 tunnel adopts 6to4 addresses. The address format is 2002:abcd:efgh:subnet number::interface ID/64, where 2002 represents the fixed IPv6 address prefix, and abcd:efgh represents the 32-bit globally unique source IPv4 address of the 6to4 tunnel, in hexadecimal

notation. For example, 1.1.1.1 can be represented by 0101:0101. The part that follows 2002:abcd:efgh uniquely identifies a host in a 6to4 network. The tunnel destination is automatically determined by the embedded IPv4 address, which makes it easy to create a 6to4 tunnel.

The tunnel can forward IPv6 packets because the 16-bit subnet number of the 64-bit address prefix in 6to4 addresses can be customized and the first 48 bits in the address prefix are fixed to a permanent value and the IPv4 address of the tunnel source or destination.

Figure 80 6to4 tunnel



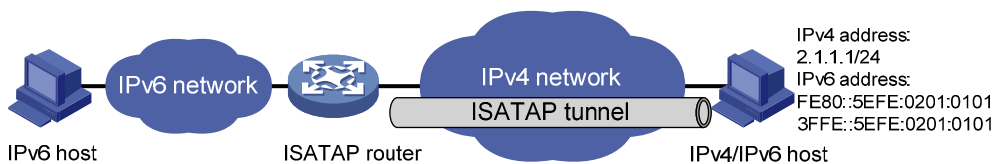
- ISATAP tunneling

An ISATAP tunnel is a point-to-multipoint automatic tunnel. The destination of a tunnel can automatically be acquired from the embedded IPv4 address in the destination address of an IPv6 packet.

When an ISATAP tunnel is used, the destination address of an IPv6 packet and the IPv6 address of a tunnel interface both adopt special ISATAP addresses. The ISATAP address format is prefix(64bit):0:5EFE:abcd:efgh. The 64-bit prefix is the prefix of a valid IPv6 unicast address, but abcd:efgh is a 32-bit source IPv4 address in hexadecimal, which might not be globally unique. Through the embedded IPv4 address, an ISATAP tunnel can be automatically created to transfer IPv6 packets.

The ISATAP tunnel is mainly used for communication between IPv6 routers or between a host and an IPv6 router over an IPv4 network.

Figure 81 Principle of ISATAP tunneling



Protocols and standards

- RFC 1853, *IP in IP Tunneling*
- RFC 2473, *Generic Packet Tunneling in IPv6 Specification*
- RFC 2893, *Transition Mechanisms for IPv6 Hosts and Routers*
- RFC 3056, *Connection of IPv6 Domains via IPv4 Clouds*
- RFC 4214, *Intra-Site Automatic Tunnel Addressing Protocol (ISATAP)*

Tunneling configuration task list

Task	Remarks
Configuring a tunnel interface	Required.
Configuring an IPv6 over IPv4 tunnel	Configuring an IPv6 manual tunnel
	Configuring a 6to4 tunnel
	Configuring an ISATAP tunnel
	Optional. Use one as needed.

Configuring a tunnel interface

Configure a Layer 3 virtual tunnel interface on each device on a tunnel so that devices at both ends can send, identify, and process packets from the tunnel.

Configuration guidelines

Follow these guidelines when you configure a tunnel interface:

- Before configuring a tunnel interface on a switch, you may need create a service loopback group with its service type as Tunnel, and add unused Layer 2 Ethernet interfaces of the switch to the service loopback group.
- On the switch, an encapsulated packet cannot be forwarded a second time at Layer 3 by using the destination address and routing table, but is sent to the loopback interface, which then sends the packet to the forwarding module for Layer 3 forwarding. You must reference a service loopback group on the tunnel interface. Otherwise, the tunnel interface will not be up and packets cannot be transmitted over the tunnel. For creation and configuration of a service loopback group, see *Layer 2—LAN Switching Configuration Guide*.
- The **tunnel bandwidth** command does not change the actual bandwidth of the tunnel interface, but sets a bandwidth value for dynamical routing protocols to calculate the cost of a tunnel path. You can determine the value according to the bandwidth of the output interface.
- The switch fragments IP unicast packets larger than the MTU on the tunnel interface, and sends an ICMP error packet to the source device to inform it to modify its MTU.
- The switch fragments IPv6 unicast packets larger than the MTU on the tunnel interface, and sends an ICMPv6 error packet to the source device to inform it to modify its MTU. The minimum MTU value in the ICMPv6 error packet is 1280 bytes, and you must configure a MTU no smaller than 1280 bytes for the tunnel (such as an IPv6 over IPv4 tunnel or an IPv6 over IPv6 tunnel).
- By default, sending of ICMP destination unreachable packets is disabled. To enable it, use the **ip unreachable enable** command.

Configuration procedure

To configure a tunnel interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a tunnel interface and enter its view.	interface tunnel <i>number</i>	By default, no tunnel interface is created.

Step	Command	Remarks
3. Configure the description for the interface.	description <i>text</i>	Optional. By default, the description of a tunnel interface is Tunnelnumber Interface .
4. Reference a service loopback group.	service-loopback-group <i>number</i>	By default, the tunnel does not reference any service loopback group.
5. Set the MTU of the tunnel interface.	mtu <i>size</i>	Optional. 64000 bytes by default. <ul style="list-style-type: none"> An MTU set on any tunnel interface is effective on all existing tunnel interfaces. You can issue this command multiple times, but only the last configuration takes effect.
6. Set the bandwidth of the tunnel interface.	tunnel bandwidth <i>bandwidth-value</i>	Optional. By default, the bandwidth of the tunnel interface is 64 kbps.
7. Restore the default setting.	default	Optional.
8. Shut down the tunnel interface.	shutdown	Optional. By default, the interface is up.

Configuring an IPv6 manual tunnel

Configuration prerequisites

Configure IP addresses for interfaces (such as the VLAN interface, and loopback interface) on the device to ensure normal communication. One of the interfaces will be used as the source interface of the tunnel.

Configuration guidelines

Follow these guidelines when you configure an IPv6 manual tunnel:

- After a tunnel interface is deleted, all the features configured on the tunnel interface will be deleted.
- To encapsulate and forward IPv6 packets whose destination address does not belong to the subnet where the current tunnel interface resides, you must configure a static route or dynamic routing for forwarding those packets through this tunnel interface. If you configure a static route to that destination IPv6 address, specify this tunnel interface as the outbound interface, or the peer tunnel interface address as the next hop. A similar configuration is required at the other tunnel end. If you configure dynamic routing at both ends, enable the dynamic routing protocol on both tunnel interfaces. For the detailed configuration, see *Layer 3—IP Routing Configuration Guide*.

Configuration procedure

To configure an IPv6 manual tunnel:

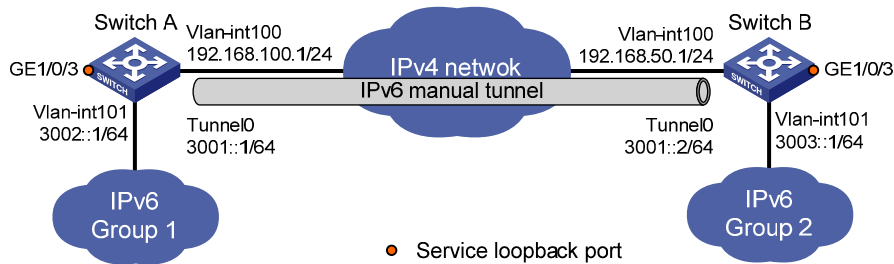
Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IPv6.	ipv6	By default, the IPv6 packet forwarding function is disabled.
3. Enter tunnel interface view.	interface tunnel <i>number</i>	N/A
4. Configure an IPv6 address for the tunnel interface.	<ul style="list-style-type: none"> • Configure a global unicast IPv6 address or a site-local address: <ul style="list-style-type: none"> ○ ipv6 address { <i>ipv6-address prefix-length</i> <i>ipv6-address/prefix-length</i> } ○ ipv6 address <i>ipv6-address/prefix-length eui-64</i> • Configure a link-local IPv6 address: <ul style="list-style-type: none"> ○ ipv6 address auto link-local ○ ipv6 address <i>ipv6-address link-local</i> 	<p>The link-local IPv6 address configuration is optional.</p> <p>By default,</p> <ul style="list-style-type: none"> • No IPv6 global unicast address or site-local address is configured for the tunnel interface. • A link-local address is automatically created when an IPv6 global unicast address or site-local address is configured.
5. Specify the IPv6 manual tunnel mode.	tunnel-protocol ipv6-ipv4	<p>IPv6 manual tunnel by default.</p> <p>The same tunnel mode should be configured at both ends of the tunnel. Otherwise, packet delivery will fail.</p>
6. Configure a source address or interface for the tunnel.	source { <i>ip-address</i> <i>interface-type interface-number</i> }	By default, no source address or interface is configured for the tunnel.
7. Configure a destination address for the tunnel.	destination <i>ip-address</i>	By default, no destination address is configured for the tunnel.
8. Return to system view.	quit	N/A
9. Enable dropping of IPv6 packets using IPv4-compatible IPv6 addresses.	tunnel discard ipv4-compatible-packet	<p>Optional.</p> <p>Disabled by default.</p>

Configuration example

Network requirements

As shown in [Figure 1](#), two IPv6 networks are connected over an IPv4 network. Configure an IPv6 over IPv4 tunnel between Switch A and Switch B to make the two IPv6 networks reachable to each other. If the destination IPv4 address cannot be automatically obtained from the destination IPv6 addresses of packets, configure an IPv6 manual tunnel.

Figure 1 Network diagram



Configuration procedure

Before configuring an IPv6 manual tunnel, make sure that Switch A and Switch B have the corresponding VLAN interfaces created and can reach each other.

- Configure Switch A:

Enable IPv6.

```
<SwitchA> system-view
[SwitchA] ipv6
```

Configure an IPv4 address for VLAN-interface 100.

```
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ip address 192.168.100.1 255.255.255.0
[SwitchA-Vlan-interface100] quit
```

Configure an IPv6 address for VLAN-interface 101.

```
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] ipv6 address 3002::1 64
[SwitchA-Vlan-interface101] quit
```

Create service loopback group 1 to support the tunnel service.

```
[SwitchA] service-loopback group 1 type tunnel
```

Assign GigabitEthernet 1/0/3 to service loopback group 1, and disable STP, NDP, and LLDP on the interface.

```
[SwitchA] interface GigabitEthernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] undo stp enable
[SwitchA-GigabitEthernet1/0/3] undo ndp enable
[SwitchA-GigabitEthernet1/0/3] undo lldp enable
[SwitchA-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchA-GigabitEthernet1/0/3] quit
```

Configure a manual IPv6 tunnel.

```
[SwitchA] interface tunnel 0
[SwitchA-Tunnel0] ipv6 address 3001::1/64
[SwitchA-Tunnel0] source vlan-interface 100
[SwitchA-Tunnel0] destination 192.168.50.1
[SwitchA-Tunnel0] tunnel-protocol ipv6-ipv4
```

Reference service loopback group 1 on the tunnel.

```
[SwitchA-Tunnel0] service-loopback-group 1
[SwitchA-Tunnel0] quit
```

Configure a static route to IPv6 Group 2 through Tunnel 0 on Switch A.

```
[SwitchA] ipv6 route-static 3003:: 64 tunnel 0
```


- **Configure Switch B**
 - # Enable IPv6.

```
<SwitchB> system-view
[SwitchB] ipv6
```
 - # Configure an IPv4 address for VLAN-interface 100.

```
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ip address 192.168.50.1 255.255.255.0
[SwitchB-Vlan-interface100] quit
```
 - # Configure an IPv6 address for VLAN-interface 101.

```
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] ipv6 address 3003::1 64
[SwitchB-Vlan-interface101] quit
```
 - # Create service loopback group 1 to support the tunnel service.

```
[SwitchB] service-loopback group 1 type tunnel
```
 - # Assign GigabitEthernet 1/0/3 to service loopback group 1, and disable STP, NDP, and LLDP.

```
[SwitchB] interface GigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] undo stp enable
[SwitchB-GigabitEthernet1/0/3] undo ndp enable
[SwitchB-GigabitEthernet1/0/3] undo lldp enable
[SwitchB-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchB-GigabitEthernet1/0/3] quit
```
 - # Configure an IPv6 manual tunnel.

```
[SwitchB] interface tunnel 0
[SwitchB-Tunnel0] ipv6 address 3001::2/64
[SwitchB-Tunnel0] source vlan-interface 100
[SwitchB-Tunnel0] destination 192.168.100.1
[SwitchB-Tunnel0] tunnel-protocol ipv6-ipv4
```
 - # Reference service loopback group 1 on the tunnel.

```
[SwitchB-Tunnel0] service-loopback-group 1
[SwitchB-Tunnel0] quit
```
 - # Configure a static route to IPv6 Group 1 through Tunnel 0 on Switch B.

```
[SwitchB] ipv6 route-static 3002:: 64 tunnel 0
```

Verifying the configuration

Display the status of the tunnel interfaces on Switch A and Switch B.

```
[SwitchA] display ipv6 interface tunnel 0
Tunnel0 current state :UP
Line protocol current state :UP
IPv6 is enabled, link-local address is FE80::C0A8:6401
Global unicast address(es):
  3001::1, subnet is 3001::/64
Joined group address(es):
  FF02::1:FF00:0
  FF02::1:FF00:1
  FF02::1:FFA8:6401
  FF02::2
  FF02::1
```

```

MTU is 1480 bytes
ND reachable time is 30000 milliseconds
ND retransmit interval is 1000 milliseconds
Hosts use stateless autoconfig for addresses
IPv6 Packet statistics:
  InReceives:                55
...
[SwitchB] display ipv6 interface tunnel 0
Tunnel0 current state :UP
Line protocol current state :UP
IPv6 is enabled, link-local address is FE80::C0A8:3201
Global unicast address(es):
  3001::2, subnet is 3001::/64
Joined group address(es):
  FF02::1:FF00:0
  FF02::1:FF00:1
  FF02::1:FFA8:3201
  FF02::2
  FF02::1
MTU is 1480 bytes
ND reachable time is 30000 milliseconds
ND retransmit interval is 1000 milliseconds
Hosts use stateless autoconfig for addresses
IPv6 Packet statistics:
  InReceives:                55
...

```

Ping the IPv6 address of VLAN-interface 101 at the peer end from Switch A.

```

[SwitchA] ping ipv6 3003::1
PING 3003::1 : 56 data bytes, press CTRL_C to break
  Reply from 3003::1
  bytes=56 Sequence=1 hop limit=64  time = 1 ms
  Reply from 3003::1
  bytes=56 Sequence=2 hop limit=64  time = 1 ms
  Reply from 3003::1
  bytes=56 Sequence=3 hop limit=64  time = 1 ms
  Reply from 3003::1
  bytes=56 Sequence=4 hop limit=64  time = 1 ms
  Reply from 3003::1
  bytes=56 Sequence=5 hop limit=64  time = 1 ms

--- 3003::1 ping statistics ---
  5 packet(s) transmitted
  5 packet(s) received
  0.00% packet loss
  round-trip min/avg/max = 1/1/1 ms

```

Configuring a 6to4 tunnel

Configuration prerequisites

Configure IP addresses for interfaces (such as the VLAN interface, and loopback interface) on the device to ensure normal communication. One of the interfaces will be used as the source interface of the tunnel.

Configuration guidelines

Follow these guidelines when you configure a 6to4 tunnel:

- No destination address needs to be configured for a 6to4 tunnel because the destination address can automatically be obtained from the IPv4 address embedded in the 6to4 IPv6 address.
- To encapsulate and forward IPv6 packets whose destination address does not belong to the subnet where the receiving tunnel interface resides, configure a static route to reach the destination IPv6 address through this tunnel interface on the device. Because automatic tunnels do not support dynamic routing, you can configure a static route to that destination IPv6 address with this tunnel interface as the outbound interface or the peer tunnel interface address as the next hop. A similar configuration is required at the other tunnel end. For the detailed configuration, see *Layer 3—IP Routing Configuration Guide*.
- Automatic tunnel interfaces using the same encapsulation protocol cannot share the same source IP address.

Configuration procedure

To configure a 6to4 tunnel:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IPv6.	ipv6	By default, the IPv6 packet forwarding function is disabled.
3. Enter tunnel interface view.	interface tunnel <i>number</i>	N/A
4. Configure an IPv6 address for the tunnel interface.	<ul style="list-style-type: none">• Configure an IPv6 global unicast address or a site-local address:<ul style="list-style-type: none">○ ipv6 address { <i>ipv6-address prefix-length</i> <i>ipv6-address/prefix-length</i> }○ ipv6 address <i>ipv6-address/prefix-length eui-64</i>• Configure an IPv6 link-local address:<ul style="list-style-type: none">○ ipv6 address auto link-local○ ipv6 address <i>ipv6-address link-local</i>	<p>The IPv6 link-local address configuration is optional.</p> <p>By default,</p> <ul style="list-style-type: none">• No IPv6 global unicast address or site-local address is configured for the tunnel interface.• A link-local address will automatically be generated when an IPv6 global unicast address or site-local address is configured.

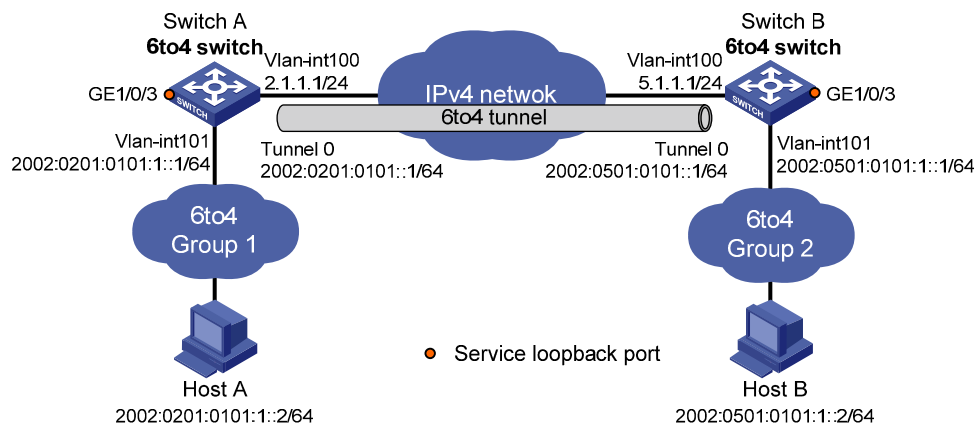
Step	Command	Remarks
5. Specify the 6to4 tunnel mode.	tunnel-protocol ipv6-ipv4 6to4	IPv6 manual tunnel by default. The same tunnel mode should be configured at both ends of the tunnel. Otherwise, packet delivery will fail.
6. Configure a source address or interface for the tunnel.	source { ip-address interface-type interface-number }	By default, no source address or interface is configured for the tunnel.
7. Return to system view.	quit	N/A
8. Enable dropping of IPv6 packets using IPv4-compatible IPv6 addresses.	tunnel discard ipv4-compatible-packet	Optional. Disabled by default.

Configuration example

Network requirements

As shown in Figure 2, two 6to4 networks are connected to an IPv4 network through two 6to4 switches (Switch A and Switch B). Configure a 6to4 tunnel to make Host A and Host B reachable to each other.

Figure 2 Network diagram



Configuration consideration

To enable communication between 6to4 networks, configure 6to4 addresses for 6to4 switches and hosts in the 6to4 networks.

- The IPv4 address of VLAN-interface 100 on Switch A is 2.1.1.1/24, and the corresponding 6to4 prefix is 2002:0201:0101::/48 after it is translated to an IPv6 address. Assign interface Tunnel 0 to subnet 2002:0201:0101::/64 and VLAN-interface 101 to subnet 2002:0201:0101:1::/64.
- The IPv4 address of VLAN-interface 100 on Switch B is 5.1.1.1/24, and the corresponding 6to4 prefix is 2002:0501:0101::/48 after it is translated to an IPv6 address. Assign interface Tunnel 0 to subnet 2002:0501:0101::/64 and VLAN-interface 101 to subnet 2002:0501:0101:1::/64.

Configuration procedure

Before configuring a 6to4 tunnel, make sure that Switch A and Switch B have the corresponding VLAN interfaces created and are reachable to each other.

- **Configure Switch A:**
 - # Enable IPv6.**

```
<SwitchA> system-view
[SwitchA] ipv6
```
 - # Configure an IPv4 address for VLAN-interface 100.**

```
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ip address 2.1.1.1 24
[SwitchA-Vlan-interface100] quit
```
 - # Configure an IPv6 address for VLAN-interface 101.**

```
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] ipv6 address 2002:0201:0101:1::1/64
[SwitchA-Vlan-interface101] quit
```
 - # Create service loopback group 1 to support the tunnel service.**

```
[SwitchA] service-loopback group 1 type tunnel
```
 - # Assign GigabitEthernet 1/0/3 to service loopback group 1, and disable STP, NDP, and LLDP on the interface.**

```
[SwitchA] interface GigabitEthernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] undo stp enable
[SwitchA-GigabitEthernet1/0/3] undo ndp enable
[SwitchA-GigabitEthernet1/0/3] undo lldp enable
[SwitchA-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchA-GigabitEthernet1/0/3] quit
```
 - # Configure a 6to4 tunnel.**

```
[SwitchA] interface tunnel 0
[SwitchA-Tunnel0] ipv6 address 2002:201:101::1/64
[SwitchA-Tunnel0] source vlan-interface 100
[SwitchA-Tunnel0] tunnel-protocol ipv6-ipv4 6to4
```
 - # Reference service loopback group 1 on the tunnel.**

```
[SwitchA-Tunnel0] service-loopback-group 1
[SwitchA-Tunnel0] quit
```
 - # Configure a static route whose destination address is 2002::/16 and next-hop is the tunnel interface.**

```
[SwitchA] ipv6 route-static 2002:: 16 tunnel 0
```
- **Configure Switch B:**
 - # Enable IPv6.**

```
<SwitchB> system-view
[SwitchB] ipv6
```
 - # Configure an IPv4 address for VLAN-interface 100.**

```
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ip address 5.1.1.1 24
[SwitchB-Vlan-interface100] quit
```
 - # Configure an IPv6 address for VLAN-interface 101.**

```
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] ipv6 address 2002:0501:0101:1::1/64
[SwitchB-Vlan-interface101] quit
```
 - # Create service loopback group 1 to support the tunnel service.**

```

[SwitchB] service-loopback group 1 type tunnel
# Assign GigabitEthernet 1/0/3 to service loopback group 1, and disable STP, NDP, and LLDP on
the interface.
[SwitchB] interface GigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] undo stp enable
[SwitchB-GigabitEthernet1/0/3] undo ndp enable
[SwitchB-GigabitEthernet1/0/3] undo lldp enable
[SwitchB-GigabitEthernet1/0/3] port service-loopback group 1
[SwitchB-GigabitEthernet1/0/3] quit
# Configure the 6to4 tunnel.
[SwitchB] interface tunnel 0
[SwitchB-Tunnel0] ipv6 address 2002:0501:0101::1/64
[SwitchB-Tunnel0] source vlan-interface 100
[SwitchB-Tunnel0] tunnel-protocol ipv6-ipv4 6to4
# Reference service loopback group 1 on the tunnel.
[SwitchB-Tunnel0] service-loopback-group 1
[SwitchB-Tunnel0] quit
# Configure a static route whose destination address is 2002::/16 and the next hop is the tunnel
interface.
[SwitchB] ipv6 route-static 2002:: 16 tunnel 0

```

Verifying the configuration

```

# Ping Host B from Host A or ping Host A from Host B.
D:\>ping6 -s 2002:201:101:1::2 2002:501:101:1::2

Pinging 2002:501:101:1::2
from 2002:201:101:1::2 with 32 bytes of data:

Reply from 2002:501:101:1::2: bytes=32 time=13ms
Reply from 2002:501:101:1::2: bytes=32 time=1ms
Reply from 2002:501:101:1::2: bytes=32 time=1ms
Reply from 2002:501:101:1::2: bytes=32 time<1ms

Ping statistics for 2002:501:101:1::2:
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
    Approximate round trip times in milli-seconds:
        Minimum = 0ms, Maximum = 13ms, Average = 3ms

```

Configuring an ISATAP tunnel

Configuration prerequisites

Configure IP addresses for interfaces (such as the VLAN interface, and loopback interface) on the device to ensure normal communication. One of the interfaces will be used as the source interface of the tunnel.

Configuration guidelines

Follow these guidelines when you configure an ISATAP tunnel:

- No destination address needs to be configured for an ISATAP tunnel. The destination address of the tunnel can be automatically obtained through the IPv4 address embedded in the ISATAP address.
- To encapsulate and forward IPv6 packets whose destination address does not belong to the subnet where the receiving tunnel interface resides, configure a static route to reach the destination IPv6 address through this tunnel interface on the device. Because automatic tunnels do not support dynamic routing, you can configure a static route to that destination IPv6 address with this tunnel interface as the outbound interface or the peer tunnel interface address as the next hop. A similar configuration is required at the other tunnel end. For the detailed configuration, see *Layer 3—IP Routing Configuration Guide*.
- The automatic tunnel interfaces using the same encapsulation protocol cannot share the same source IP address.

Configuration procedure

To configure an ISATAP tunnel:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IPv6.	ipv6	By default, the IPv6 forwarding function is disabled.
3. Enter tunnel interface view.	interface tunnel number	N/A
4. Configure an IPv6 address for the tunnel interface.	<ul style="list-style-type: none"> • Configure an IPv6 global unicast address or site-local address: <ul style="list-style-type: none"> ◦ ipv6 address { <i>ipv6-address prefix-length</i> <i>ipv6-address/prefix-length</i> } ◦ ipv6 address <i>ipv6-address/prefix-length eui-64</i> • Configure an IPv6 link-local address: <ul style="list-style-type: none"> ◦ ipv6 address auto link-local ◦ ipv6 address <i>ipv6-address link-local</i> 	<p>The IPv6 link-local address configuration is optional.</p> <p>By default,</p> <ul style="list-style-type: none"> • No IPv6 global unicast address is configured for the tunnel interface. • A link-local address will automatically be generated when an IPv6 global unicast address or link-local address is configured.
5. Specify the ISATAP tunnel mode.	tunnel-protocol ipv6-ipv4 isatap	<p>IPv6 manual tunnel by default.</p> <p>The same tunnel mode should be configured at both ends of the tunnel. Otherwise, packet delivery will fail.</p>
6. Configure a source address or interface for the tunnel.	source { <i>ip-address</i> <i>interface-type interface-number</i> }	By default, no source address or interface is configured for the tunnel.
7. Return to system view.	quit	N/A

Step	Command	Remarks	
8.	Enable dropping of IPv6 packets using IPv4-compatible IPv6 addresses.	tunnel discard ipv4-compatible-packet	Optional. Disabled by default.

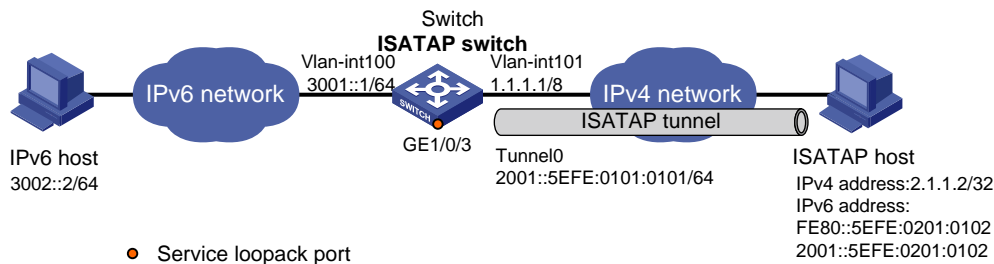
Configuration example

Network requirements

As shown in [Figure 3](#), an IPv6 network is connected to an IPv4 network through an ISATAP switch. IPv6 hosts reside in the IPv4 network.

Configure the IPv6 hosts to access the IPv6 network through the ISATAP tunnel.

Figure 3 Network diagram



Configuration procedure

Before configuring an ISATAP tunnel, make sure that the corresponding VLAN interfaces have been created on the switch, and that VLAN-interface 101 on the ISATAP switch and the ISATAP host can reach each other.

- Configure the switch:


```
# Enable IPv6.
<Switch> system-view
[Switch] ipv6

# Configure addresses for interfaces.
[Switch] interface vlan-interface 100
[Switch-Vlan-interface100] ipv6 address 3001::1/64
[Switch-Vlan-interface100] quit
[Switch] interface vlan-interface 101
[Switch-Vlan-interface101] ip address 1.1.1.1 255.0.0.0
[Switch-Vlan-interface101] quit

# Create service loopback group 1 to support the tunnel service.
[Switch] service-loopback group 1 type tunnel

# Assign GigabitEthernet 1/0/3 to service loopback group 1, and disable STP, NDP, and LLDP on the interface.
[Switch] interface GigabitEthernet 1/0/3
[Switch-GigabitEthernet1/0/3] undo stp enable
[Switch-GigabitEthernet1/0/3] undo ndp enable
[Switch-GigabitEthernet1/0/3] undo lldp enable
```



```
[Switch-GigabitEthernet1/0/3] port service-loopback group 1
[Switch-GigabitEthernet1/0/3] quit
```

Configure an ISATAP tunnel.

```
[Switch] interface tunnel 0
[Switch-Tunnel0] ipv6 address 2001::5efe:0101:0101 64
[Switch-Tunnel0] source vlan-interface 101
[Switch-Tunnel0] tunnel-protocol ipv6-ipv4 isatap
```

Disable the RA suppression so that hosts can acquire information such as the address prefix from the RA message released by the ISATAP switch.

```
[Switch-Tunnel0] undo ipv6 nd ra halt
```

Reference service loopback group 1 on the tunnel.

```
[Switch-Tunnel0] service-loopback-group 1
[Switch-Tunnel0] quit
```

Configure a static route to the ISATAP host.

```
[Switch] ipv6 route-static 2001:: 16 tunnel 0
```

- Configure the ISATAP host:

The specific configuration on the ISATAP host is related to its operating system. The following example shows the configuration of the host running the Windows XP.

Install IPv6.

```
C:\>ipv6 install
```

On a Windows XP-based host, the ISATAP interface is usually interface 2. Configure the IPv4 address of the ISATAP router on the interface to complete the configuration on the host. Before doing that, display the ISATAP interface information:

```
C:\>ipv6 if 2
```

```
Interface 2: Automatic Tunneling Pseudo-Interface
  Guid {48FCE3FC-EC30-E50E-F1A7-71172AEEE3AE}
  does not use Neighbor Discovery
  does not use Router Discovery
  routing preference 1
  EUI-64 embedded IPv4 address: 0.0.0.0
  router link-layer address: 0.0.0.0
  preferred link-local fe80::5efe:2.1.1.2, life infinite
  link MTU 1280 (true link MTU 65515)
  current hop limit 128
  reachable time 42500ms (base 30000ms)
  retransmission interval 1000ms
  DAD transmits 0
  default site prefix length 48
```

A link-local address (fe80::5efe:2.1.1.2) in the ISATAP format was automatically generated for the ISATAP interface. Configure the IPv4 address of the ISATAP switch on the ISATAP interface.

```
C:\>ipv6 rlu 2 1.1.1.1
```

Display the information on the ISATAP interface.

```
C:\>ipv6 if 2
```

```
Interface 2: Automatic Tunneling Pseudo-Interface
  Guid {48FCE3FC-EC30-E50E-F1A7-71172AEEE3AE}
  does not use Neighbor Discovery
  uses Router Discovery
```

```

routing preference 1
EUI-64 embedded IPv4 address: 2.1.1.2
router link-layer address: 1.1.1.1
  preferred global 2001::5efe:2.1.1.2, life 29d23h59m46s/6d23h59m46s (public)
  preferred link-local fe80::5efe:2.1.1.2, life infinite
link MTU 1500 (true link MTU 65515)
current hop limit 255
reachable time 42500ms (base 30000ms)
retransmission interval 1000ms
DAD transmits 0
default site prefix length 48

```

By comparison, it is found that the host acquires the address prefix 2001::/64 and automatically generates the address 2001::5efe:2.1.1.2. Meanwhile, "uses Router Discovery" is displayed, indicating that the router discovery function is enabled on the host. At this time, ping the IPv6 address of the tunnel interface of the switch. If the address is successfully pinged, an ISATAP tunnel is established.

```
C:\>ping 2001::5efe:1.1.1.1
```

```
Pinging 2001::5efe:1.1.1.1 with 32 bytes of data:
```

```

Reply from 2001::5efe:1.1.1.1: time=1ms
Reply from 2001::5efe:1.1.1.1: time=1ms
Reply from 2001::5efe:1.1.1.1: time=1ms
Reply from 2001::5efe:1.1.1.1: time=1ms

```

```

Ping statistics for 2001::5efe:1.1.1.1:
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
    Approximate round trip times in milli-seconds:
        Minimum = 1ms, Maximum = 1ms, Average = 1ms

```

Verifying the configuration

The ISATAP host can access the host in the IPV6 network.

Displaying and maintaining tunneling configuration

Task	Command	Remarks
Display information about tunnel interfaces.	<pre> display interface [tunnel] [brief [down]] [{ begin exclude include } <i>regular-expression</i>] display interface tunnel <i>number</i> [brief] [{ begin exclude include } <i>regular-expression</i>] </pre>	Available in any view
Display IPv6 information on tunnel interfaces.	<pre> display ipv6 interface tunnel [<i>number</i>] [brief] [{ begin exclude include } <i>regular-expression</i>] </pre>	Available in any view
Clear statistics on tunnel interfaces.	<pre> reset counters interface [tunnel [<i>number</i>]] </pre>	Available in user view

Troubleshooting tunneling configuration

Symptom

After the configuration of related parameters such as tunnel source address, tunnel destination address, and tunnel mode, the tunnel interface is still not up.

Solution

1. The common cause is that the physical interface of the tunnel source is not up. Use the **display interface tunnel** or **display ipv6 interface tunnel** commands to view whether the physical interface of the tunnel source is up. If the physical interface is down, check the network connections.
2. Another possible cause is that the tunnel destination is unreachable. Use the **display ipv6 routing-table** or **display ip routing-table** command to view whether the tunnel destination is reachable. If no routing entry is available for tunnel communication in the routing table, configure related routes.

Index

[A](#) [B](#) [C](#) [D](#) [E](#) [I](#) [O](#) [P](#) [S](#) [T](#) [U](#)

A

- Address/prefix lease renewal, [144](#)
- Application environment of trusted ports, [72](#)
- Applying an extended address pool on an interface, [47](#)
- Applying the address pool to an interface, [150](#)
- ARP configuration examples, [7](#)
- Assigning an IP address to an interface, [24](#)

B

- BOOTP client configuration example, [84](#)

C

- Configuration guidelines, [12](#)
- Configuration procedure, [109](#)
- Configuration procedure, [12](#)
- Configuration procedure, [99](#)
- Configuration procedure, [21](#)
- Configuration restrictions, [83](#)
- Configuration restrictions, [68](#)
- Configuration restrictions and guidelines, [109](#)
- Configuring a 6to4 tunnel, [185](#)
- Configuring a DHCPv6 address pool, [149](#)
- Configuring a DHCPv6 snooping trusted port, [165](#)
- Configuring a static ARP entry, [3](#)
- Configuring a tunnel interface, [179](#)
- Configuring an address pool for the DHCP server, [38](#)
- Configuring an interface to dynamically obtain an IP address through BOOTP, [84](#)
- Configuring an IPv6 manual tunnel, [180](#)
- Configuring an ISATAP tunnel, [188](#)
- Configuring ARP quick update, [5](#)
- Configuring basic IPv6 functions, [121](#)
- Configuring DHCP packet rate limit, [79](#)
- Configuring DHCP snooping basic functions, [75](#)
- Configuring DHCP snooping entries backup, [77](#)
- Configuring DHCP snooping to support Option 82, [76](#)
- Configuring DNS spoofing, [90](#)
- Configuring ICMP to send error packets, [106](#)

- Configuring ICMPv6 packet sending, [134](#)
- Configuring IP unnumbered(only available on the HP 5500 EI), [26](#)
- Configuring IPv6 ND, [125](#)
- Configuring IPv6 TCP properties, [133](#)
- Configuring multicast ARP, [6](#)
- Configuring path MTU discovery, [133](#)
- Configuring stateless DHCPv6, [145](#)
- Configuring TCP attributes, [104](#)
- Configuring the DHCP relay agent security functions, [60](#)
- Configuring the DHCP relay agent to release an IP address, [62](#)
- Configuring the DHCP relay agent to support Option 82, [63](#)
- Configuring the DHCP server security functions, [48](#)
- Configuring the DHCPv6 client, [160](#)
- Configuring the DHCPv6 relay agent, [156](#)
- Configuring the DNS proxy, [89](#)
- Configuring the IPv4 DNS client, [88](#)
- Configuring the IPv6 DNS client, [168](#)
- Configuring the maximum number of DHCPv6 snooping entries an interface can learn, [166](#)
- Configuring the maximum number of dynamic ARP entries for an interface, [4](#)
- Configuring trusted ports in a cascaded network, [72](#)
- Correlating a DHCP server group with a relay agent interface, [59](#)
- Creating a prefix pool, [149](#)

D

- DHCP address allocation, [28](#)
- DHCP client configuration example, [69](#)
- DHCP message format, [30](#)
- DHCP options, [31](#)
- DHCP relay agent configuration examples, [65](#)
- DHCP relay agent configuration task list, [58](#)
- DHCP server configuration examples, [51](#)
- DHCP server configuration task list, [38](#)
- DHCP snooping configuration examples, [80](#)

- DHCP snooping configuration task list, [74](#)
- DHCP snooping functions, [71](#)
- DHCP snooping support for Option 82, [73](#)
- DHCPv6 address/prefix assignment, [143](#)
- DHCPv6 relay agent configuration example, [157](#)
- DHCPv6 server configuration example, [151](#)
- DHCPv6 server configuration task list, [148](#)
- DHCPv6 snooping configuration example, [166](#)
- Displaying and maintaining ARP, [6](#)
- Displaying and maintaining ARP snooping, [21](#)
- Displaying and maintaining BOOTP client configuration, [84](#)
- Displaying and maintaining DHCP snooping, [80](#)
- Displaying and maintaining DHCPv6 snooping, [166](#)
- Displaying and maintaining IP addressing, [27](#)
- Displaying and maintaining IP performance optimization, [108](#)
- Displaying and maintaining IPv4 DNS, [91](#)
- Displaying and maintaining IPv6 basics configuration, [136](#)
- Displaying and maintaining IPv6 DNS, [169](#)
- Displaying and maintaining proxy ARP, [15](#)
- Displaying and maintaining the DHCP client, [69](#)
- Displaying and maintaining the DHCP relay agent, [64](#)
- Displaying and maintaining the DHCP server, [50](#)
- Displaying and maintaining the DHCPv6 client, [161](#)
- Displaying and maintaining the DHCPv6 relay agent, [157](#)
- Displaying and maintaining the DHCPv6 server, [151](#)
- Displaying and maintaining tunneling configuration, [192](#)
- Displaying and maintaining UDP helper, [110](#)
- DNS proxy configuration example, [95](#)
- Dynamic domain name resolution configuration example, [171](#)
- Dynamic domain name resolution configuration example, [92](#)

E

- Enabling client offline detection, [49](#)
- Enabling common proxy ARP, [14](#)
- Enabling DHCP, [46](#)
- Enabling DHCP, [58](#)
- Enabling DHCP starvation attack protection, [78](#)
- Enabling DHCP-REQUEST message attack protection, [79](#)

- Enabling DHCPv6 snooping, [165](#)
- Enabling dynamic ARP entry check, [5](#)
- Enabling handling of Option 82, [49](#)
- Enabling local proxy ARP, [14](#)
- Enabling offline detection, [62](#)
- Enabling receiving and forwarding of directed broadcasts to a directly connected network, [103](#)
- Enabling the DHCP client on an interface, [68](#)
- Enabling the DHCP relay agent on an interface, [58](#)
- Enabling the DHCP server on an interface, [46](#)
- Enabling the DHCPv6 server, [149](#)

I

- Introduction to DHCPv6, [143](#)
- IPv6 basics configuration example, [137](#)
- IPv6 basics configuration task list, [120](#)
- IRDP configuration example, [100](#)

O

- Overview(Configuring IPv4 DNS), [85](#)
- Overview(Configuring BOOTP client), [83](#)
- Overview(Configuring DHCP relay agent), [56](#)
- Overview(Configuring DHCPv6 relay agent), [155](#)
- Overview(Configuring DHCPv6 client), [160](#)
- Overview(Configuring UDP helper), [109](#)
- Overview(Configuring IPv6 basics), [112](#)
- Overview(Configuring ARP), [1](#)
- Overview(Configuring gratuitous ARP), [11](#)
- Overview(Configuring proxy ARP), [13](#)
- Overview(Configuring DHCPv6 server), [147](#)
- Overview(Configuring IRDP), [98](#)
- Overview(Configuring ARP snooping), [21](#)
- Overview(Configuring IP addressing), [22](#)
- Overview(Configuring IPv6 DNS), [168](#)
- Overview(Configuring DHCP server), [36](#)
- Overview(Configuring tunneling), [176](#)
- Overview(Configuring DHCPv6 snooping), [164](#)

P

- Protocols and standards, [146](#)
- Protocols and standards, [35](#)
- Proxy ARP configuration examples, [15](#)

S

- Setting the aging timer for dynamic ARP entries, [4](#)
- Setting the DSCP value for DHCP packets, [64](#)
- Setting the DSCP value for DHCP packets, [68](#)

Setting the DSCP value for DHCP packets, [50](#)
Setting the DSCP value for DHCPv6 packets, [151](#)
Setting the DSCP value for DHCPv6 packets, [157](#)
Setting the DSCP value for DHCPv6 packets, [160](#)
Setting the DSCP value for DNS packets, [90](#)
Setting the DSCP value for IPv6 DNS packets, [169](#)
Specifying the source interface for DNS packets, [90](#)
Specifying the threshold for sending trap messages, [50](#)
Stateless DHCPv6 configuration example, [161](#)
Static domain name resolution configuration example, [91](#)

Static domain name resolution configuration example, [170](#)

T

Troubleshooting DHCP relay agent configuration, [67](#)
Troubleshooting DHCP server configuration, [55](#)
Troubleshooting IPv4 DNS configuration, [97](#)
Troubleshooting IPv6 basics configuration, [142](#)
Troubleshooting tunneling configuration, [193](#)
Tunneling configuration task list, [178](#)

U

UDP helper configuration example, [110](#)

Contents

IP routing basics	1
Hardware compatibility.....	1
Overview.....	1
Routing table.....	1
Dynamic routing protocols	2
Routing preference.....	3
Load sharing.....	3
Route backup.....	4
Route recursion.....	4
Route redistribution	4
Displaying and maintaining a routing table	4
Configuring static routing	6
Hardware compatibility.....	6
Introduction	6
Static route	6
Default route.....	6
Static route configuration items.....	6
Configuring a static route.....	7
Configuring BFD for static routes.....	8
BFD control packet mode	8
BFD echo packet mode.....	9
Configuring static route FRR.....	10
Configuration prerequisites	10
Configuration guidelines	10
Configuration procedure	10
Displaying and maintaining static routes	11
Static route configuration examples	11
Basic static route configuration example	11
Static route FRR configuration example	13
BFD for static routes configuration example (direct session)	15
BFD for static routes configuration example (indirect session)	17
Configuring RIP	21
Hardware compatibility.....	21
Overview.....	21
Understanding RIP	21
RIP operation.....	22
RIP versions	22
RIP message format	23
Supported RIP features	24
Protocols and standards	25
RIP configuration task list.....	25
Configuring RIP basic functions	26
Configuration prerequisites	26
Enabling RIP and a RIP interface	26
Configuring the interface behavior	26
Configuring a RIP version.....	27
Setting the DSCP value for RIP packets	28
Configuring RIP route control.....	28

Configuring an additional routing metric	28
Configuring RIPv2 route summarization	29
Disabling host route reception	29
Advertising a default route	30
Configuring inbound or outbound route filtering	30
Configuring a priority for RIP	31
Configuring RIP route redistribution	31
Tuning and optimizing RIP networks	32
Configuring RIP timers	32
Configuring split horizon and poison reverse	32
Configuring the maximum number of ECMP routes	33
Enabling zero field check on incoming RIPv1 messages	33
Enabling source IP address check on incoming RIP updates	34
Configuring RIPv2 message authentication	34
Specifying a RIP neighbor	34
Configuring RIP-to-MIB binding	35
Configuring the RIP packet sending rate	35
Configuring RIP FRR	36
Configuring BFD for RIP	36
Single-hop detection in BFD echo packet mode	37
Bidirectional detection in BFD control packet mode	37
Displaying and maintaining RIP	37
RIP configuration examples	38
Configuring RIP version	38
Configuring RIP route redistribution	39
Configuring an additional metric for a RIP interface	42
Configuring RIP to advertise a summary route	43
RIP FRR configuration example	46
Configuring BFD for RIP (single-hop detection in BFD echo packet mode)	47
Configuring BFD for RIP (bidirectional detection in BFD control packet mode)	50
Troubleshooting RIP	54
No RIP updates received	54
Route oscillation occurred	54
Configuring OSPF	55
Hardware compatibility	55
Introduction to OSPF	55
Basic concepts	55
Area based OSPF network partition	57
Router types	60
OSPF network classification	61
DR and BDR	62
OSPF packet formats	63
Supported features	71
Protocols and standards	72
OSPF configuration task list	72
Enabling OSPF	73
Configuration prerequisites	73
Configuration procedure	74
Configuring OSPF areas	75
Configuration prerequisites	75
Configuring a stub area	75
Configuring an NSSA area	76
Configuring a virtual link	76
Configuring OSPF network types	77

Configuration prerequisites	77
Configuring the OSPF network type for an interface as broadcast	77
Configuring the OSPF network type for an interface as NBMA	78
Configuring the OSPF network type for an interface as P2MP	78
Configuring the OSPF network type for an interface as P2P	79
Configuring OSPF route control	79
Configuration prerequisites	79
Configuring OSPF route summarization	80
Configuring OSPF inbound route filtering	81
Configuring ABR Type-3 LSA filtering	81
Configuring an OSPF cost for an interface	81
Configuring the maximum number of OSPF routes	82
Configuring the maximum number of ECMP routes	82
Configuring OSPF preference	83
Configuring OSPF route redistribution	83
Advertising a host route	85
Tuning and optimizing OSPF networks	85
Configuration prerequisites	85
Configuring OSPF packet timers	85
Specifying LSA transmission delay	86
Specifying SPF calculation interval	87
Specifying the LSA arrival interval	87
Specifying the LSA generation interval	87
Disabling interfaces from receiving and sending OSPF packets	88
Configuring stub routers	88
Configuring OSPF authentication	89
Adding the interface MTU into DD packets	89
Configuring the maximum number of external LSAs in LSDB	90
Enabling compatibility with RFC 1583	90
Logging neighbor state changes	90
Configuring OSPF network management	91
Enabling message logging	91
Enabling the advertisement and reception of opaque LSAs	91
Configuring OSPF to give priority to receiving and processing hello packets	92
Configuring the LSU transmit rate	92
Setting the DSCP value for OSPF packets	93
Enabling OSPF ISPF	93
Configuring OSPF FRR	93
Configuring OSPF Graceful Restart	95
Configuring the OSPF GR Restarter	95
Configuring the OSPF GR Helper	96
Triggering OSPF Graceful Restart	96
Configuring BFD for OSPF	97
Configuring control packet bidirectional detection	97
Configuring echo packet single-hop detection	97
Displaying and maintaining OSPF	97
OSPF configuration examples	99
Configuring OSPF basic functions	99
Configuring OSPF route redistribution	102
Configuring OSPF to advertise a summary route	103
Configuring an OSPF stub area	106
Configuring an OSPF NSSA area	108
Configuring OSPF DR election	110
Configuring OSPF virtual links	114

Configuring OSPF Graceful Restart	116
Configuring route filtering	119
Configuring OSPF FRR	121
Configuring BFD for OSPF	123
Troubleshooting OSPF configuration	127
No OSPF neighbor relationship established	127
Incorrect routing information	128
Configuring IS-IS	129
Hardware compatibility	129
IS-IS overview	129
Basic concepts	129
IS-IS area	131
IS-IS network type	133
IS-IS PDU format	134
Supported IS-IS features	140
Protocols and standards	142
IS-IS configuration task list	142
Configuring IS-IS basic functions	143
Configuration prerequisites	143
Enabling IS-IS	143
Configuring the IS level and circuit level	144
Configuring the network type of an interface as P2P	144
Configuring IS-IS routing information control	145
Configuration prerequisites	145
Configuring IS-IS link cost	145
Specifying a priority for IS-IS	146
Configuring the maximum number of ECMP routes	147
Configuring IS-IS route summarization	147
Advertising a default route	147
Configuring IS-IS route redistribution	148
Configuring IS-IS route filtering	148
Configuring IS-IS route leaking	149
Tuning and optimizing IS-IS networks	150
Configuration prerequisites	150
Specifying intervals for sending IS-IS hello and CSNP packets	150
Specifying the IS-IS hello multiplier	150
Configuring a DIS priority for an interface	151
Disabling an interface from sending or receiving IS-IS packets	151
Enabling an interface to send small hello packets	151
Configuring LSP parameters	152
Configuring SPF parameters	154
Assigning a high priority to IS-IS routes	154
Setting the LSDB overload bit	155
Configuring system ID to host name mappings	155
Enabling the logging of neighbor state changes	156
Enhancing IS-IS network security	157
Configuration prerequisites	157
Configuring neighbor relationship authentication	157
Configuring area authentication	157
Configuring routing domain authentication	158
Configuring IS-IS GR	158
Configuring IS-IS NSR	159
Configuring IS-IS FRR	159
Enabling IS-IS SNMP trap	160

Binding an IS-IS process with MIBs	161
Configuring BFD for IS-IS	161
Displaying and maintaining IS-IS	161
IS-IS configuration examples	162
IS-IS basic configuration	162
DIS election configuration	167
Configuring IS-IS route redistribution	171
IS-IS Graceful Restart configuration example	174
IS-IS NSR configuration example	176
IS-IS FRR configuration example	178
IS-IS authentication configuration example	181
Configuring BFD for IS-IS	183
Configuring BGP	187
Hardware compatibility	187
BGP overview	187
BGP messages formats	187
BGP path attributes	190
BGP route selection	194
IBGP and IGP synchronization	196
Settlements for problems in large scale BGP networks	196
BGP GR	199
MP-BGP	200
Protocols and standards	200
BGP configuration task list	201
Configuring BGP basic functions	202
Configuration prerequisites	202
Creating a BGP connection	202
Specifying the source interface for TCP connections	203
Allowing establishment of EBGP connection to an indirectly connected peer or peer group	204
Controlling route generation	204
Configuration prerequisites	204
Injecting a local network	204
Configuring BGP route redistribution	205
Enabling default route redistribution into BGP	205
Controlling route distribution and reception	206
Configuration prerequisites	206
Configuring BGP route summarization	206
Advertising a default route to a peer or peer group	206
Configuring BGP route distribution/reception filtering policies	207
Enabling BGP and IGP route synchronization	209
Limiting prefixes received from a peer or peer group	209
Configuring BGP route dampening	210
Configuring a shortcut route	210
Configuring BGP route attributes	211
Configuration prerequisites	211
Specifying a preferred value for routes received	211
Configuring preferences for BGP routes	211
Configuring the default local preference	212
Configuring the MED attribute	212
Configuring the next hop attribute	214
Configuring the AS-PATH attribute	215
Tuning and optimizing BGP networks	217
Configuration prerequisites	217
Configuring the BGP keepalive interval and holdtime	217

Configuring the interval for sending the same update.....	218
Configuring BGP soft-reset.....	218
Enabling the BGP ORF capability.....	219
Enabling 4-byte AS number suppression.....	220
Setting the DSCP value for BGP packets.....	221
Enabling quick EBGP session reestablishment.....	221
Enabling MD5 authentication for TCP connections.....	222
Configuring BGP load balancing.....	222
Forbidding session establishment with a peer or peer group.....	222
Configuring a large scale BGP network.....	223
Configuration prerequisites.....	223
Configuring BGP peer groups.....	223
Configuring BGP community.....	225
Configuring a BGP route reflector.....	225
Configuring a BGP confederation.....	226
Configuring BGP GR.....	227
Enabling trap.....	227
Enabling logging of peer state changes.....	228
Configuring BFD for BGP.....	228
Displaying and maintaining BGP.....	228
Displaying BGP.....	228
Resetting BGP connections.....	230
Clearing BGP information.....	230
BGP configuration examples.....	230
BGP basic configuration.....	230
BGP and IGP synchronization configuration.....	234
BGP load balancing configuration.....	237
BGP community configuration.....	240
BGP route reflector configuration.....	242
BGP confederation configuration.....	244
BGP path selection configuration.....	247
BGP GR configuration.....	250
Configuring BFD for BGP.....	251
Troubleshooting BGP.....	256
BGP peer relationship not established.....	256
Configuring IPv6 static routing.....	257
Hardware compatibility.....	257
Overview.....	257
IPv6 static routes features.....	257
Default IPv6 route.....	257
Configuring an IPv6 static route.....	257
Displaying and maintaining IPv6 static routes.....	258
IPv6 static routing configuration example.....	258
Configuring RIPng.....	261
Hardware compatibility.....	261
Introduction to RIPng.....	261
RIPng working mechanism.....	261
RIPng packet format.....	262
RIPng packet processing procedure.....	263
Protocols and standards.....	263
RIPng configuration task list.....	263
Configuring RIPng basic functions.....	264
Configuring RIPng route control.....	264

Configuring an additional routing metric	264
Configuring RIPng route summarization	265
Advertising a default route	265
Configuring a RIPng route filtering policy	265
Configuring a priority for RIPng	266
Configuring RIPng route redistribution	266
Tuning and optimizing the RIPng network	266
Configuring RIPng timers	266
Configuring split horizon and poison reverse	267
Configuring zero field check on RIPng packets	268
Configuring the maximum number of ECMP routes	268
Applying IPsec policies for RIPng	268
Configuration prerequisites	268
Configuration guidelines	269
Configuration procedure	269
Displaying and maintaining RIPng	269
RIPng configuration examples	270
Configuring RIPng basic functions	270
Configuring RIPng route redistribution	272
Configuring RIPng IPsec policies	275
Configuring OSPFv3	278
Hardware compatibility	278
Introduction to OSPFv3	278
OSPFv3 overview	278
OSPFv3 packets	278
OSPFv3 LSA types	279
OSPFv3 timers	279
OSPFv3 features supported	280
Protocols and standards	280
OSPFv3 configuration task list	280
Enabling OSPFv3	281
Configuration prerequisites	281
Enabling OSPFv3	281
Configuring OSPFv3 area parameters	282
Configuration prerequisites	282
Configuring an OSPFv3 stub area	282
Configuring an OSPFv3 virtual link	283
Configuring OSPFv3 network types	283
Configuration prerequisites	283
Configuring the OSPFv3 network type for an interface	283
Configuring an NBMA or P2MP neighbor	284
Configuring OSPFv3 routing information control	284
Configuration prerequisites	284
Configuring OSPFv3 route summarization	284
Configuring OSPFv3 inbound route filtering	285
Configuring an OSPFv3 cost for an interface	285
Configuring the maximum number of OSPFv3 ECMP routes	286
Configuring a priority for OSPFv3	286
Configuring OSPFv3 route redistribution	286
Tuning and optimizing OSPFv3 networks	287
Configuration prerequisites	287
Configuring OSPFv3 timers	287
Configuring a DR priority for an interface	288
Ignoring MTU check for DD packets	289

Disabling interfaces from receiving and sending OSPFv3 packets	289
Enabling the logging of neighbor state changes	289
Configuring OSPFv3 GR	290
Configuring GR Restarter	290
Configuring GR Helper	290
Configuring BFD for OSPFv3	291
Applying IPsec policies for OSPFv3	291
Displaying and maintaining OSPFv3	293
OSPFv3 configuration examples	294
Configuring OSPFv3 areas	294
Configuring OSPFv3 DR election	297
Configuring OSPFv3 route redistribution	300
Configuring OSPFv3 GR	303
Configuring BFD for OSPFv3	305
Configuring OSPFv3 IPsec policies	308
Troubleshooting OSPFv3 configuration	311
No OSPFv3 neighbor relationship established	311
Incorrect routing information	312
Configuring IPv6 IS-IS	313
Hardware compatibility	313
Introduction to IPv6 IS-IS	313
Configuring IPv6 IS-IS basic functions	313
Configuring IPv6 IS-IS routing information control	314
Configuring BFD for IPv6 IS-IS	315
Displaying and maintaining IPv6 IS-IS	315
IPv6 IS-IS configuration examples	316
IPv6 IS-IS basic configuration example	316
BFD for IPv6 IS-IS configuration example	321
Configuring IPv6 BGP	325
Hardware compatibility	325
IPv6 BGP overview	325
IPv6 BGP configuration task list	325
Configuring IPv6 BGP basic functions	327
Configuration prerequisites	327
Configuration guidelines	327
Specifying an IPv6 BGP peer	327
Injecting a local IPv6 route	327
Configuring a preferred value for routes from a peer or peer group	328
Specifying the source interface for establishing TCP connections	328
Allowing the establishment of an indirect EBGP connection	329
Configuring a description for an IPv6 peer or peer group	329
Disabling session establishment to an IPv6 peer or peer group	329
Logging IPv6 peer or peer group state changes	330
Controlling route distribution and reception	330
Configuration prerequisites	330
Configuring IPv6 BGP route redistribution	330
Configuring IPv6 BGP route summarization	331
Advertising a default route to an IPv6 peer or peer group	331
Configuring outbound route filtering	332
Configuring inbound route filtering	332
Configuring IPv6 BGP and IGP route synchronization	333
Configuring route dampening	334
Configuring IPv6 BGP route attributes	334

Configuration prerequisites	334
Configuring IPv6 BGP preference and default LOCAL_PREF and NEXT_HOP attributes	334
Configuring the MED attribute	335
Configuring the AS_PATH attribute	335
Tuning and optimizing IPv6 BGP networks	336
Configuration prerequisites	337
Configuring IPv6 BGP timers	337
Configuring IPv6 BGP soft reset	337
Enabling the IPv6 BGP ORF capability	338
Enabling 4-byte AS number suppression	339
Setting the DSCP value for IPv6 BGP packets	340
Configuring the maximum number of load-balanced routes	340
Enabling MD5 authentication for TCP connections	340
Applying an IPsec policy to an IPv6 BGP peer or peer group	341
Configuring a large-scale IPv6 BGP network	342
Configuration prerequisites	342
Configuring IPv6 BGP peer group	342
Configuring IPv6 BGP community	343
Configuring an IPv6 BGP route reflector	344
Configuring BFD for IPv6 BGP	345
Displaying and maintaining IPv6 BGP	345
Displaying BGP	345
Resetting IPv6 BGP connections	346
Clearing IPv6 BGP information	347
IPv6 BGP configuration examples	347
IPv6 BGP basic configuration example	347
IPv6 BGP route reflector configuration example	349
IPv6 BGP IPsec policy configuration example	350
Configuring BFD for IPv6 BGP	355
Troubleshooting IPv6 BGP configuration	360
IPv6 BGP peer relationship not established	360
Configuring routing policies	361
Hardware compatibility	361
Introduction to routing policy	361
Routing policy application	361
Routing policy implementation	361
Filters	361
Defining filters	363
Configuration prerequisites	363
Defining an IP-prefix list	363
Defining an AS path list	364
Defining a community list	364
Defining an extended community list	365
Configuring a routing policy	365
Configuration prerequisites	365
Creating a routing policy	365
Defining if-match clauses	366
Defining apply clauses	367
Defining a continue clause	368
Displaying and maintaining the routing policy	369
Routing policy configuration examples	370
Applying a routing policy to IPv4 route redistribution	370
Applying a routing policy to IPv6 route redistribution	372
Applying a routing policy to filter received BGP routes	374

Troubleshooting routing policy configuration	376
IPv4 routing information filtering failure.....	376
IPv6 routing information filtering failure.....	377
Configuring policy-based routing	377
Hardware compatibility.....	377
Introduction to PBR.....	377
PBR modes.....	378
Concepts.....	378
QoS mode.....	379
Configuring PBR (using a PBR policy).....	379
Defining a policy.....	379
Configuring local PBR.....	380
Configuring interface PBR.....	380
PBR and track.....	381
Configuring PBR (using a QoS policy).....	381
Configuring a QoS policy.....	381
Applying the QoS policy.....	382
Displaying and maintaining PBR configuration.....	383
PBR configuration (using a PBR policy).....	383
PBR configuration (using a QoS policy).....	383
PBR configuration examples.....	384
Configuring local PBR based on packet type.....	384
Configuring interface PBR based on packet type.....	385
IPv4 PBR configuration example (using a QoS policy).....	387
IPv6 PBR configuration example (using a QoS policy).....	388
Configuring MCE	390
MCE overview.....	390
MPLS L3VPN overview.....	390
MPLS L3VPN concepts.....	391
Multi-VPN-instance CE.....	393
How MCE works.....	393
Configuring routing on an MCE.....	394
Route exchange between an MCE and a VPN site.....	394
Route exchange between an MCE and a PE.....	395
Configuring an MCE.....	396
Configuring VPN instances.....	396
Configuring routing on an MCE.....	398
Configuration prerequisites.....	398
Configuring routing between MCE and VPN site.....	398
Configuring routing between MCE and PE.....	403
Resetting BGP connections.....	407
Displaying and maintaining MCE.....	408
MCE configuration examples.....	409
Using OSPF to advertise VPN routes to the PE.....	409
Using BGP to advertise VPN routes to the PE.....	414
Configuring IPv6 MCE.....	418
Overview.....	418
Configuring an IPv6 MCE.....	418
Configuring VPN instances.....	418
Configuring routing on an IPv6 MCE.....	420
Configuration prerequisites.....	420
Configuring routing between IPv6 MCE and VPN site.....	420
Configuring routing between IPv6 MCE and PE.....	424

Resetting BGP connections	427
Displaying information about IPv6 MCE	427
IPv6 MCE configuration examples	428
Using IPv6 ISIS to advertise VPN routes to the PE	428
Index	435

IP routing basics

Hardware compatibility

The HP 5500 SI Switch Series does not support configuring Layer 3 Ethernet interfaces.

The HP 5500 SI Switch Series does not support VPN-related parameters.

The HP 5500 SI Switch Series does not support OSPF, BGP, IS-IS, OSPFv3, IPv6 BGP, or IPv6 IS-IS.

Overview

IP routing directs the forwarding of IP packets on routers based on a routing table. This book focuses on unicast routing protocols. For more information about multicast routing protocols, see *IP Multicast Configuration Guide*.

The term "router" in this chapter refers to both routers and Layer 3 switches.

The types of interfaces that appear in any figures other than the network diagrams for configuration examples are for illustration only. Some of them might be unavailable on your switch.

The term "interface" in the routing features refers to Layer 3 interfaces, including VLAN interfaces and route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).

Routing table

A router maintains at least two routing tables: one global routing table and one forwarding information base (FIB). The FIB table contains only the optimal routes, and the global routing table contains all routes. The router uses the FIB table to forward packets. For more information about the FIB table, see *Layer 3—IP Services Configuration Guide*.

Routes can be classified by different criteria, as shown in [Table 1](#).

Table 1 Categories of routes

Criterion	Categories
Destination	<ul style="list-style-type: none">• Network route—Destination is a network. The subnet mask is less than 32 bits.• Host route—Destination is a host. The subnet mask is 32 bits.
Whether the destination is directly connected	<ul style="list-style-type: none">• Direct route—Destination is directly connected.• Indirect route—Destination is indirectly connected.
Origin	<ul style="list-style-type: none">• Direct route—A direct route is discovered by the data link protocol on an interface, and is also called an "interface route."• Static route—A static route is manually configured by an administrator.• Dynamic route—A dynamic route is dynamically discovered by a routing protocol.

Static routes are easy to configure and require less system resources. They work well in small and stable networks. In networks where topology changes may occur frequently, using a dynamic routing protocol is better.

To display brief information about a routing table, use the **display ip routing-table** command, as shown in the following example:

```
<Sysname> display ip routing-table
Routing Tables: Public
          Destinations : 7          Routes : 7

Destination/Mask    Proto  Pre  Cost           NextHop         Interface
1.1.1.0/24          Direct  0    0             1.1.1.1         Vlan11
2.2.2.0/24          Static  60   0             12.2.2.2        Vlan12
80.1.1.0/24         OSPF   10   2             80.1.1.1        Vlan13
...
```

A route entry includes the following key items:

- **Destination**—IP address of the destination host or network.
- **Mask**—The network mask specifies, in company with the destination address, the address of the destination network. A logical AND operation between the destination address and the network mask yields the address of the destination network. For example, if the destination address is 129.102.8.10 and the mask 255.255.0.0, the address of the destination network is 129.102.0.0. A network mask is made up of a certain number of consecutive 1s. It can be expressed in dotted decimal format or by the number of the 1s.
- **Pre**—Preference of the route. Among routes to the same destination, the one with the highest preference is optimal.
- **Cost**—When multiple routes to a destination have the same preference, the one with the smallest cost becomes the optimal route.
- **NextHop**—Specifies the IP address of the next hop.
- **Interface**—Specifies the interface through which a matching IP packet is to be forwarded.

Dynamic routing protocols

Dynamic routing protocols dynamically collect and report reachability information to adapt to topology changes. They are suitable for large networks.

Compared with static routing, dynamic routing protocols require more resources, and are complicated to configure.

Dynamic routing protocols can be classified based on different criteria, as shown in [Table 2](#):

Table 2 Dynamic routing protocols

Criterion	Categories
Optional scope	<ul style="list-style-type: none"> • Interior gateway protocols (IGPs)—Work within an autonomous system (AS). Examples include RIP, OSPF, and IS-IS. • Exterior gateway protocols (EGPs)—Work between ASs. The most popular one is BGP.

Criterion	Categories
Routing algorithm	<ul style="list-style-type: none"> • Distance-vector protocols—RIP and BGP. BGP is also considered a path-vector protocol. • Link-state protocols—OSPF and IS-IS.
Destination address type	<ul style="list-style-type: none"> • Unicast routing protocols—RIP, OSPF, BGP, and IS-IS. • Multicast routing protocols—PIM-SM and PIM-DM (For more information, see <i>IP Multicast Configuration Guide</i>).
IP version	<ul style="list-style-type: none"> • IPv4 routing protocols—RIP, OSPF, BGP, and IS-IS. • IPv6 routing protocols—RIPng, OSPFv3, IPv6 BGP, and IPv6 IS-IS.

NOTE:

An AS refers to a group of routers sharing the same routing policy and working under the same administration.

Routing preference

Different routing protocols can find different routes to the same destination. However, not all of those routes are optimal. For route selection, routing protocols, direct routes, and static routes are assigned different preferences. The route with the highest preference is preferred.

The preference of a direct route is always 0 and cannot be changed. You can manually configure preferences for any other route type. Each static route can be configured with a different preference. The following table lists the types of routes and the default preferences. The smaller the preference value, the higher the preference.

Table 3 Route types and their default route preferences

Routing approach	Preference
Direct route	0
OSPF	10
IS-IS	15
Static route	60
RIP	100
OSPF ASE	150
OSPF NSSA	150
IBGP	255
EBGP	255
Unknown (route from an untrusted source)	256

Load sharing

A routing protocol can be configured with multiple equal-cost routes to the same destination. These routes have the same preference and will all be used to accomplish load sharing if there is no route with a higher preference available.

Static routing/IPv6 static routing, RIP/RIPng, OSPF/OSPFv3, BGP/IPv6 BGP, and IS-IS/IPv6 IS-IS supports ECMP load sharing.

The load sharing function is available only on the HP 5500 EI Switch series.

Route backup

Route backup can improve network availability. Among multiple routes to the same destination, the route with the highest preference is the main route and all others are backup routes.

The router forwards matching packets through the main route. When the main route fails, the route with the highest preference among the backup routes is selected to forward packets. When the main route recovers, the router uses it to forward packets.

Route recursion

To use a BGP, static, or RIP route that has an indirectly-connected next hop, a router must perform route recursion to find the outgoing interface to reach the next hop.

Link-state routing protocols, such as OSPF and IS-IS, do not need route recursion, because they obtain directly-connected next hops through route calculation.

Route redistribution

Route redistribution enables routing protocols to learn route information from each other. A dynamic routing protocol can redistribute routes from other protocols including direct routes and static routes. For more information, see the respective chapters on those routing protocols in this configuration guide.

Displaying and maintaining a routing table

Task	Command	Remarks
Display information about the routing table.	display ip routing-table [vpn-instance <i>vpn-instance-name</i>] [verbose] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display information about routes permitted by an IPv4 basic ACL.	display ip routing-table [vpn-instance <i>vpn-instance-name</i>] acl <i>acl-number</i> [verbose] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display information about routes to the specified destination.	display ip routing-table [vpn-instance <i>vpn-instance-name</i>] <i>ip-address</i> [<i>mask</i> <i>mask-length</i>] [longer-match] [verbose] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display information about routes with destination addresses in the specified range.	display ip routing-table [vpn-instance <i>vpn-instance-name</i>] <i>ip-address1</i> { <i>mask</i> <i>mask-length</i> } <i>ip-address2</i> { <i>mask</i> <i>mask-length</i> } [verbose] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view

Task	Command	Remarks
Display routing information permitted by an IPv4 prefix list.	display ip routing-table [vpn-instance <i>vpn-instance-name</i>] ip-prefix <i>ip-prefix-name</i> [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display routes of a routing protocol.	display ip routing-table [vpn-instance <i>vpn-instance-name</i>] protocol <i>protocol</i> [inactive verbose] [{ begin exclude include } <i>regular-expression</i>] [{ begin exclude include } <i>regular-expression</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display statistics about the routing table.	display ip routing-table [vpn-instance <i>vpn-instance-name</i>] statistics [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear statistics for the routing table.	reset ip routing-table statistics protocol [vpn-instance <i>vpn-instance-name</i>] { <i>protocol</i> all }	Available in user view
Display IPv6 routing table information.	display ipv6 routing-table [vpn-instance <i>vpn-instance-name</i>] [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display routing information permitted by an IPv6 ACL.	display ipv6 routing-table [vpn-instance <i>vpn-instance-name</i>] acl <i>acl6-number</i> [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display routing information for a specified destination IPv6 address.	display ipv6 routing-table [vpn-instance <i>vpn-instance-name</i>] <i>ipv6-address</i> <i>prefix-length</i> [longer-match] [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 routing information for an IPv6 address range.	display ipv6 routing-table [vpn-instance <i>vpn-instance-name</i>] <i>ipv6-address1</i> <i>prefix-length1</i> <i>ipv6-address2</i> <i>prefix-length2</i> [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display routing information permitted by an IPv6 prefix list.	display ipv6 routing-table [vpn-instance <i>vpn-instance-name</i>] ipv6-prefix <i>ipv6-prefix-name</i> [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 routing information of a routing protocol.	display ipv6 routing-table [vpn-instance <i>vpn-instance-name</i>] protocol <i>protocol</i> [inactive verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 routing statistics.	display ipv6 routing-table [vpn-instance <i>vpn-instance-name</i>] statistics [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear specified IPv6 routing statistics.	reset ipv6 routing-table statistics protocol [vpn-instance <i>vpn-instance-name</i>] { <i>protocol</i> all }	Available in user view

Configuring static routing

Hardware compatibility

The HP 5500 SI Switch Series does not support VPN and BFD related parameters or FRR.

Introduction

Static route

Static routes are manually configured. If a network's topology is simple, you only need to configure static routes for the network to work properly. The proper configuration and usage of static routes can improve network performance and ensure bandwidth for important network applications.

The disadvantage of using static routes is that they cannot adapt to network topology changes. If a fault or a topological change occurs in the network, the relevant routes will be unreachable and the network breaks. When this happens, the network administrator must modify the static routes manually.

The term "router" in this chapter refers to both routers and Layer 3 switches.

Default route

Without a default route, a packet that does not match any routing entries is discarded.

A default route is used to forward packets that do not match any routing entry. It can be configured in either of the following ways:

- The network administrator can configure a default route with both the destination and mask being 0.0.0.0. The router forwards any packet whose destination address fails to match any entry in the routing table to the next hop of the default static route.
- Some dynamic routing protocols, such as OSPF, RIP, and IS-IS, can also generate a default route. For example, an upstream router running OSPF can generate a default route and advertise it to other routers, which install the default route with the next hop being the upstream router.

Static route configuration items

Before you configure a static route, you must know the following concepts:

- Destination address and mask
In the **ip route-static** command, an IPv4 address is in dotted decimal format. A mask can be either in dotted decimal format or in the form of mask length—the number of consecutive 1s in the mask.
- Output interface and next hop address
When you configure a static route, specify either the output interface, next hop address, or both depending on the specific occasion. The next hop address cannot be a local interface IP address; otherwise, the route configuration will not take effect.

Each route lookup operation has to find the next hop to resolve the destination link layer address.

When you specify the output interface, follow these guidelines:

- If the output interface is a Null 0 interface, no next hop address is required.
- If you specify a broadcast interface (such as an Ethernet interface or VLAN interface) as the output interface, you must specify the corresponding next hop for the output interface.
- Other attributes

You can configure different priorities for different static routes so that route management policies can be more flexible. For example, specifying the same priority for different routes to the same destination enables load sharing, but specifying different priorities for these routes enables route backup.

NOTE:

The load sharing function is available only on the HP 5500 EI Switch series.

Configuring a static route

Before you configure a static route, complete the following tasks:

- Configure the physical parameters for related interfaces.
- Configure the link-layer attributes for related interfaces.
- Configure the IP addresses for related interfaces.

Follow these guidelines when you configure a static route:

- The next hop address cannot be the IP address of a local interface (such as Ethernet interface and VLAN interface). Otherwise, the static route does not take effect.
- If you do not specify the preference when you configure a static route, the default preference will be used. Reconfiguring the default preference applies only to newly created static routes.
- You can flexibly control static routes by configuring tag values and using the tag values in the routing policy.
- If the destination IP address and mask are both configured as 0.0.0.0 with the **ip route-static** command, then the route is the default route.
- For more information about track, see *High Availability Configuration Guide*.

To configure a static route:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Configure a static route.	<ul style="list-style-type: none"> Approach 1: ip route-static <i>dest-address</i> { <i>mask</i> <i>mask-length</i> } { <i>next-hop-address</i> [track <i>track-entry-number</i>] <i>interface-type</i> <i>interface-number</i> [<i>next-hop-address</i>] vpn-instance <i>d-vpn-instance-name</i> <i>next-hop-address</i> [track <i>track-entry-number</i>] } [preference <i>preference-value</i>] [tag <i>tag-value</i>] [permanent] [description <i>description-text</i>] Approach 2: ip route-static vpn-instance <i>s-vpn-instance-name</i>&<1-6> <i>dest-address</i> { <i>mask</i> <i>mask-length</i> } { <i>next-hop-address</i> [public] [track <i>track-entry-number</i>] <i>interface-type</i> <i>interface-number</i> [<i>next-hop-address</i>] vpn-instance <i>d-vpn-instance-name</i> <i>next-hop-address</i> [track <i>track-entry-number</i>] } [preference <i>preference-value</i>] [tag <i>tag-value</i>] [permanent] [description <i>description-text</i>] 	<p>Use either approach.</p> <p>By default, preference for static routes is 60, tag is 0, and no description information is configured.</p> <p>Do not specify the permanent and track keywords simultaneously.</p> <p>If the output interface is down, the permanent static route is still active.</p>
3. Configure the default preference for static routes.	ip route-static default-preference <i>default-preference-value</i>	Optional. 60 by default.

Configuring BFD for static routes

NOTE:

Only the HP 5500 EI Switch Series supports BFD for static routes.

Bidirectional forwarding detection (BFD) provides a general-purpose, standard, medium-, and protocol-independent fast failure detection mechanism. It can uniformly and quickly detect the failures of the bidirectional forwarding paths between two routers for protocols, such as routing protocols. For more information about BFD, see *High Availability Configuration Guide*.

A dynamic routing protocol notifies BFD of its neighbor information. BFD uses such information to establish sessions with neighbors by sending BFD control packets. Static routing has no neighbor discovery mechanism. This section describes how static routing implements BFD.

BFD control packet mode

To use BFD control packets for bidirectional detection between two devices, you must enable BFD control packet mode for each device's static route destined to the peer.

To configure a static route and enable BFD control packet mode for it, specify an outbound interface and a direct next hop—BFD establishes a direct session, or specify an indirect next hop and a specific BFD packet source address—BFD establishes an indirect session—for the static route.

To configure a static route with BFD control packet mode enabled (direct session):

Step	Command	Remarks
4. Enter system view.	system-view	N/A
5. Configure a static route and enable BFD control packet mode for it.	<ul style="list-style-type: none"> Approach 1: ip route-static <i>dest-address</i> { <i>mask</i> <i>mask-length</i> } <i>interface-type interface-number next-hop-address</i> bfd control-packet [preference <i>preference-value</i>] [tag <i>tag-value</i>] [description <i>description-text</i>] Approach 2: ip route-static vpn-instance <i>s-vpn-instance-name</i>&<1-6> <i>dest-address</i> { <i>mask</i> <i>mask-length</i> } <i>interface-type</i> <i>interface-number next-hop-address</i> bfd control-packet [preference <i>preference-value</i>] [tag <i>tag-value</i>] [description <i>description-text</i>] 	Use either approach.

To configure a static route with BFD control packet mode enabled (indirect session):

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure a static route and enable BFD control packet mode for it.	<ul style="list-style-type: none"> Approach 1: ip route-static <i>dest-address</i> { <i>mask</i> <i>mask-length</i> } <i>next-hop-address</i> bfd control-packet bfd-source <i>ip-address</i> [preference <i>preference-value</i>] [tag <i>tag-value</i>] [description <i>description-text</i>] Approach 2: ip route-static vpn-instance <i>s-vpn-instance-name</i>&<1-6> <i>dest-address</i> { <i>mask</i> <i>mask-length</i> } <i>next-hop-address</i> bfd control-packet bfd-source <i>ip-address</i> [preference <i>preference-value</i>] [tag <i>tag-value</i>] [description <i>description-text</i>] 	Use either command.

BFD echo packet mode

ⓘ IMPORTANT:

- Enabling BFD for a flapping route could worsen the situation.
- Do not use BFD for a static route with the outbound interface having the spoofing attribute.

With BFD echo packet mode enabled for a static route, the local device sends BFD echo packets to the peer, which loops it back to test the link. For the echo mode, only one end needs to establish the BFD session, and the source address of echo packets must be configured.

To configure BFD echo packet mode for static routes:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the source address of echo packets.	bfd echo-source-ip <i>ip-address</i>	Not configured by default.

Step	Command	Remarks
3. Enable BFD echo packet mode for static routes.	<ul style="list-style-type: none"> Approach 1: <code>ip route-static dest-address { mask mask-length } interface-type interface-number next-hop-address bfd echo-packet [preference preference-value] [tag tag-value] [description description-text]</code> Approach 2: <code>ip route-static vpn-instance s-vpn-instance-name&<1-6> dest-address { mask mask-length } interface-type interface-number next-hop-address bfd echo-packet [preference preference-value] [tag tag-value] [description description-text]</code> 	Use either approach.

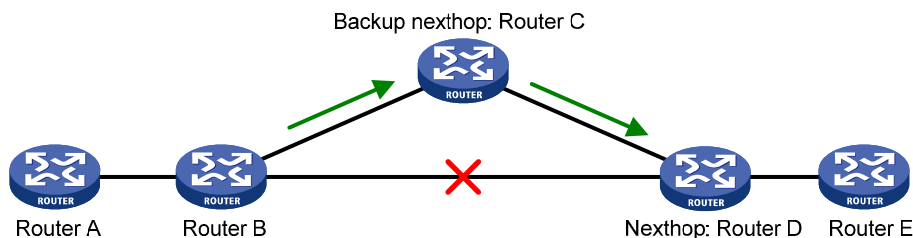
Configuring static route FRR

NOTE:

Only the HP 5500 EI Switch Series supports static route FRR.

When a link or a router fails, the packets on the path may be discarded, or a routing loop occurs. To avoid such problems, you can enable static route fast reroute (FRR).

Figure 1 Network diagram



As shown in Figure 1, upon a link failure, FRR designates a backup next hop by using a routing policy for routes matching the specified criteria. Packets are directed to the backup next hop to avoid traffic interruption.

Configuration prerequisites

Configuring static route FRR needs to reference a routing policy. You can specify a backup next hop in a routing policy by using the **apply fast-reroute backup-interface** command. For more information about the command and routing policy configurations, see "[Configuring routing policies.](#)"

Configuration guidelines

- FRR takes effect only for static routes that have both an outbound interface and next hop.
- Do not use FRR and BFD at the same time.

Configuration procedure

To configure static route FRR:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the source address of echo packets.	bfd echo-source-ip <i>ip-address</i>	Not configured by default.
3. Configure static route FRR.	ip route-static [vpn-instance <i>vpn-instance-name</i>] fast-reroute route-policy <i>route-policy-name</i>	Not configured by default.

Displaying and maintaining static routes

Task	Command	Remarks
Display information of static routes.	display ip routing-table protocol static [inactive verbose] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Delete all the static routes.	delete [vpn-instance <i>vpn-instance-name</i>] static-routes all	Available in system view

For more information about the **display ip routing-table protocol static** [**inactive** | **verbose**] [[{ **begin** | **exclude** | **include** } *regular-expression*] command, see *Layer 3—IP Routing Command Reference*.

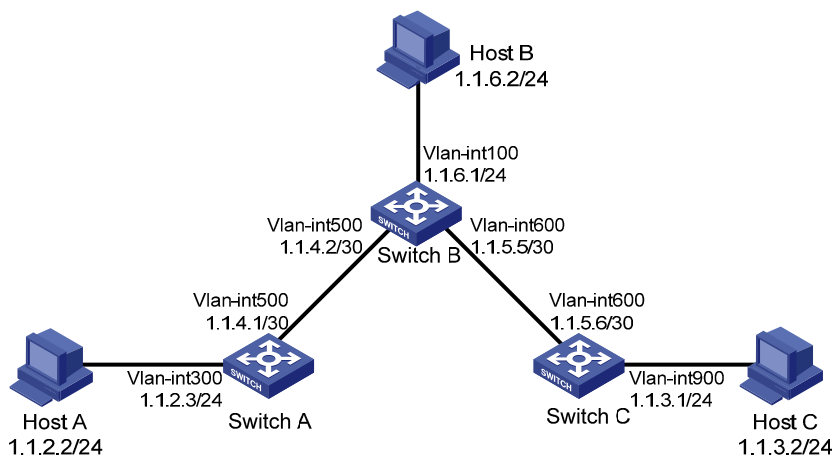
Static route configuration examples

Basic static route configuration example

Network requirements

Configure static routes in Figure 2 for interconnection between any two hosts.

Figure 2 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure static routes:

Configure a default route on Switch A.

```
<SwitchA> system-view
[SwitchA] ip route-static 0.0.0.0 0.0.0.0 1.1.4.2
```

Configure two static routes on Switch B.

```
<SwitchB> system-view
[SwitchB] ip route-static 1.1.2.0 255.255.255.0 1.1.4.1
[SwitchB] ip route-static 1.1.3.0 255.255.255.0 1.1.5.6
```

Configure a default route on Switch C.

```
<SwitchC> system-view
[SwitchC] ip route-static 0.0.0.0 0.0.0.0 1.1.5.5
```

3. Configure the hosts:

Configure the default gateways of hosts A, B, and C as 1.1.2.3, 1.1.6.1, and 1.1.3.1. (Details not shown.)

4. Display the configuration:

Display the IP routing table on Switch A.

```
[SwitchA] display ip routing-table
```

Routing Tables: Public

Destinations : 7 Routes : 7

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
0.0.0.0/0	Static	60	0	1.1.4.2	Vlan500
1.1.2.0/24	Direct	0	0	1.1.2.3	Vlan300
1.1.2.3/32	Direct	0	0	127.0.0.1	InLoop0
1.1.4.0/30	Direct	0	0	1.1.4.1	Vlan500
1.1.4.1/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0

Display the IP routing table on Switch B.

```
[SwitchB] display ip routing-table
```

Routing Tables: Public

Destinations : 10 Routes : 10

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
1.1.2.0/24	Static	60	0	1.1.4.1	Vlan500
1.1.3.0/24	Static	60	0	1.1.5.6	Vlan600
1.1.4.0/30	Direct	0	0	1.1.4.2	Vlan500
1.1.4.2/32	Direct	0	0	127.0.0.1	InLoop0
1.1.5.4/30	Direct	0	0	1.1.5.5	Vlan600
1.1.5.5/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0
1.1.6.0/24	Direct	0	0	1.1.6.1	Vlan100
1.1.6.1/32	Direct	0	0	127.0.0.1	InLoop0

Use the **ping** command on Host B to test the reachability of Host A (assuming Windows XP runs on the two hosts).

```
C:\Documents and Settings\Administrator>ping 1.1.2.2
```

```
Pinging 1.1.2.2 with 32 bytes of data:
```

```
Reply from 1.1.2.2: bytes=32 time=1ms TTL=255
```

```
Reply from 1.1.2.2: bytes=32 time=1ms TTL=255
```

```
Reply from 1.1.2.2: bytes=32 time=1ms TTL=255
```

```
Reply from 1.1.2.2: bytes=32 time=1ms TTL=255
```

```
Ping statistics for 1.1.2.2:
```

```
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
```

```
Approximate round trip times in milli-seconds:
```

```
    Minimum = 1ms, Maximum = 1ms, Average = 1ms
```

Use the **tracert** command on Host B to test the reachability of Host A.

```
C:\Documents and Settings\Administrator>tracert 1.1.2.2
```

```
Tracing route to 1.1.2.2 over a maximum of 30 hops
```

```
  0  <1 ms    <1 ms    <1 ms    1.1.6.1
  1  <1 ms    <1 ms    <1 ms    1.1.4.1
  2  1 ms     <1 ms    <1 ms    1.1.2.2
```

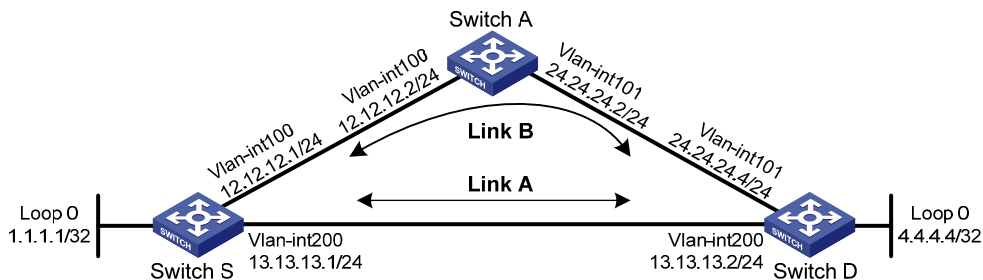
```
Trace complete.
```

Static route FRR configuration example

Network requirements

Switch S, Switch A, and Switch D are interconnected through static routes, as illustrated in [Figure 3](#). Configure static route FRR so that when the link between Switch S and Switch D fails, traffic can be switched to Link B immediately.

Figure 3 Network diagram



Configuration procedure

1. Configure IP addresses for the interfaces on each switch and configure static routes:

Follow [Figure 3](#) to configure the IP address and subnet mask of each interface on the switches. (Details not shown.)

Configure static routes on Switch S, Switch A, and Switch D so that Switch S can reach Loopback 0 on Switch D and Switch D can reach Loopback 0 on Switch S.

Configure a static route on Switch S.

```
<SwitchS> system-view
[SwitchS] ip route-static 4.4.4.4 32 vlan-interface 200 13.13.13.2
[SwitchS] ip route-static 4.4.4.4 32 vlan-interface 100 12.12.12.2 preference 65
```

Configure a static route on Switch D.

```
<SwitchD> system-view
[SwitchD] ip route-static 1.1.1.1 32 vlan-interface 200 13.13.13.1
[SwitchD] ip route-static 1.1.1.1 32 vlan-interface 101 24.24.24.2 preference 65
```

Configure a static route on Switch A.

```
<SwitchA> system-view
[SwitchA] ip route-static 4.4.4.4 32 vlan-interface 101 24.24.24.4
[SwitchA] ip route-static 1.1.1.1 32 vlan-interface 100 12.12.12.1
```

2. Configure static route FRR:

Configure Switch S.

```
[SwitchS] bfd echo-source-ip 1.1.1.1
[SwitchS] ip ip-prefix abc index 10 permit 4.4.4.4 32
[SwitchS] route-policy frr permit node 10
[SwitchS-route-policy] if-match ip-prefix abc
[SwitchS-route-policy] apply fast-reroute backup-interface vlan-interface 100
backup-nexthop 12.12.12.2
[SwitchS-route-policy] quit
[SwitchS] ip route-static fast-reroute route-policy frr
```

Configure Switch D.

```
[SwitchD] bfd echo-source-ip 4.4.4.4
[SwitchD] ip ip-prefix abc index 10 permit 1.1.1.1 32
[SwitchD] route-policy frr permit node 10
[SwitchD-route-policy] if-match ip-prefix abc
[SwitchD-route-policy] apply fast-reroute backup-interface vlan-interface 101
backup-nexthop 24.24.24.2
[SwitchD-route-policy] quit
[SwitchD] ip route-static fast-reroute route-policy frr
```

3. Verify the configuration:

Display route 4.4.4.4/32 on Switch S to view the backup next hop information.

```
[SwitchS] display ip routing-table 4.4.4.4 verbose
```

```
Routing Table : Public
```

```
Summary Count : 1
```

```

Destination: 4.4.4.4/32
  Protocol: Static          Process ID: 0
  Preference: 60           Cost: 0
  IpPrecedence:           QoSLeId:
    NextHop: 13.13.13.2    Interface: vlan 200
    BkNextHop: 12.12.12.2 BkInterface: vlan 100
    RelyNextHop: 0.0.0.0   Neighbor : 0.0.0.0
    Tunnel ID: 0x0         Label: NULL
    BKTunnel ID: 0x0      BKLabel: NULL
    State: Active Adv     Age: 00h01m27s
```

```

Tag: 0
# Display route 1.1.1.1/32 on Switch D to view the backup next hop information.
[SwitchD] display ip routing-table 1.1.1.1 verbose
Routing Table : Public
Summary Count : 1

Destination: 1.1.1.1/32
  Protocol: Static          Process ID: 0
  Preference: 60           Cost: 0
  IpPrecedence:           QoSLeId:
    NextHop: 13.13.13.1    Interface: vlan 200
    BkNextHop: 24.24.24.2  BkInterface: vlan 101
  RelyNextHop: 0.0.0.0     Neighbor : 0.0.0.0
  Tunnel ID: 0x0           Label: NULL
  BKTunnel ID: 0x0        BKLabel: NULL
  State: Active Adv       Age: 00h01m27s
  Tag: 0

```

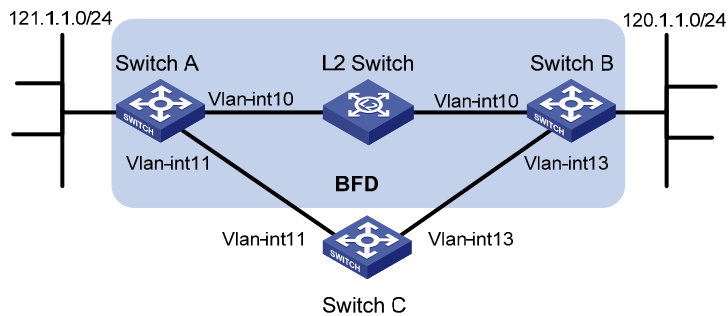
BFD for static routes configuration example (direct session)

Network requirements

As shown in [Figure 4](#), configure static routes to subnet 120.1.1.0/24 on Switch A, static routes to subnet 121.1.1.0/24 on Switch B, and static routes to subnets 120.1.1.0/24 and 121.1.1.0/24 on Switch C.

Enable BFD so that when the link between Switch A and Switch B through the Layer 2 switch fails, BFD can detect the failure immediately and Switch A and Switch B can communicate through Switch C.

Figure 4 Network diagram



Device	Interface	IP address	Device	Interface	IP address
Switch A	Vlan-int10	12.1.1.1/24	Switch B	Vlan-int10	12.1.1.2/24
	Vlan-int11	10.1.1.102/24		Vlan-int13	13.1.1.1/24
Switch C	Vlan-int11	10.1.1.100/24			
	Vlan-int13	13.1.1.2/24			

Configuration procedure

1. Configure IP addresses for the interfaces. (Details not shown.)
2. Configure BFD:

```

# Configure static routes on Switch A and enable BFD control packet mode for the static route
through the Layer 2 switch.

```



```

<SwitchA> system-view
[SwitchA] interface vlan-interface10
[SwitchA-vlan-interface10] bfd min-transmit-interval 500
[SwitchA-vlan-interface10] bfd min-receive-interval 500
[SwitchA-vlan-interface10] bfd detect-multiplier 9
[SwitchA-vlan-interface10] quit
[SwitchA] ip route-static 120.1.1.0 24 vlan-interface 10 12.1.1.2 bfd control-packet
[SwitchA] ip route-static 120.1.1.0 24 vlan-interface 11 10.1.1.100 preference 65
[SwitchA] quit

```

Configure static routes on Switch B and enable BFD control packet mode for the static route through the Layer 2 switch.

```

<SwitchB> system-view
[SwitchB] interface vlan-interface10
[SwitchB-vlan-interface10] bfd min-transmit-interval 500
[SwitchB-vlan-interface10] bfd min-receive-interval 500
[SwitchB-vlan-interface10] bfd detect-multiplier 9
[SwitchB-vlan-interface10]] quit
[SwitchB] ip route-static 121.1.1.0 24 vlan-interface 10 12.1.1.1 bfd control-packet
[SwitchB] ip route-static 121.1.1.0 24 vlan-interface 13 13.1.1.2 preference 65
[SwitchB] quit

```

Configure static routes on Switch C.

```

<SwitchC> system-view
[SwitchC] ip route-static 120.1.1.0 24 vlan-interface 13 13.1.1.1
[SwitchC] ip route-static 121.1.1.0 24 vlan-interface 11 10.1.1.102

```

3. Verify the configuration:

The following operations are performed on Switch A. The operations on Switch B are similar.

Display BFD sessions on Router A.

```

<SwitchA> display bfd session

```

```

Total Session Num: 1          Init Mode: Active

```

```

Session Working Under Ctrl Mode:

```

LD/RD	SourceAddr	DestAddr	State	Holdtime	Interface
4/7	12.1.1.1	12.1.1.2	Up	2000ms	Vlan10

Display static routes on Switch A.

```

<SwitchA> display ip routing-table protocol static

```

```

Public Routing Table : Static

```

```

Summary Count : 2

```

```

Static Routing table Status : < Active>

```

```

Summary Count : 1

```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
120.1.1.0/24	Static	60	0	12.1.1.2	Vlan10

```

Direct Routing table Status : <Inactive>

```

```

Summary Count : 1

```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
120.1.1.0/24	Static	65	0	10.1.1.100	Vlan11

Enable BFD debugging on Switch A. When the link between Switch A and Layer 2 switch fails, Switch A can detect the failure.

```
<SwitchA> debugging bfd event
```

```
<SwitchA> debugging bfd scm
```

```
<SwitchA> terminal debugging
```

```
%Jul 27 10:18:18:672 2007 SwitchA BFD/4/LOG:Sess[12.1.1.1/12.1.1.2, Vlan10,Ctrl], Sta: UP->DOWN, Diag: 1
```

```
*Jul 27 10:18:18:672 2007 SwitchA BFD/7/EVENT:Send sess-down Msg, [Src:12.1.1.1,Dst:12.1.1.2, Vlan10,Ctrl], instance:0, protocol:STATIC
```

```
*Jul 27 10:18:19:172 2007 SwitchA BFD/7/EVENT:Receive Delete-sess, [Src:12.1.1.1,Dst:12.1.1.2, Vlan10,Ctrl], Direct, Instance:0x0, Proto:STATIC
```

```
*Jul 27 10:18:19:172 2007 SwitchA BFD/7/EVENT:Notify driver to stop receiving bfd control packet
```

Display the static route information on Switch A again. Switch A communicates with Switch B over the static route passing Switch C now.

```
<SwitchA> display ip routing-table protocol static
```

```
Public Routing Table : Static
```

```
Summary Count : 2
```

```
Static Routing table Status : < Active>
```

```
Summary Count : 1
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
120.1.1.0/24	Static	65	0	10.1.1.100	Vlan11

```
Static Routing table Status : < Inactive>
```

```
Summary Count : 1
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
120.1.1.0/24	Static	60	0	12.1.1.2	Vlan10

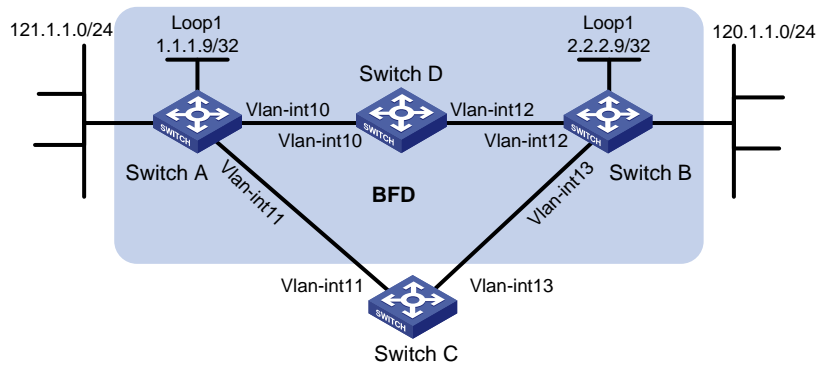
BFD for static routes configuration example (indirect session)

Network requirements

As shown in [Figure 5](#), Switch A has a route to interface Loopback1 (2.2.2.9/32) on Switch B, with outbound interface VLAN-interface 10. Switch B has a route to interface Loopback1 (1.1.1.9/32) on Switch A, with outbound interface VLAN-interface 12. Switch D has a route to 1.1.1.9/32, with outbound interface VLAN-interface 10, and a route to 2.2.2.9/32, with outbound interface VLAN-interface 12.

Configure static routes to subnet 120.1.1.0/24 on Switch A, static routes to subnet 121.1.1.0/24 on Switch B, and static routes to subnets 120.1.1.0/24 and 121.1.1.0/24 on both Switch C and Switch D. Enable BFD so that when the link between Switch A and Switch B through Switch D fails, BFD can detect the failure immediately and Switch A and Switch B can communicate through Switch C.

Figure 5 Network diagram



Device	Interface	IP address	Device	Interface	IP address
Switch A	Vlan-int10	12.1.1.1/24	Switch B	Vlan-int12	11.1.1.1/24
	Vlan-int11	10.1.1.102/24		Vlan-int13	13.1.1.1/24
	Loop1	1.1.1.9/32		Loop1	2.2.2.9/32
Switch C	Vlan-int11	10.1.1.100/24	Switch D	Vlan-int10	12.1.1.2/24
	Vlan-int13	13.1.1.2/24		Vlan-int12	11.1.1.2/24

Configuration procedure

1. Configure IP addresses for the interfaces. (Details not shown.)
2. Configure BFD:

Configure static routes on Switch A and enable BFD control packet mode for the static route through Switch D.

```
<SwitchA> system-view
[SwitchA] interface loopback 1
[SwitchA-LoopBack1] bfd min-transmit-interval 500
[SwitchA-LoopBack1] bfd min-receive-interval 500
[SwitchA-LoopBack1] bfd detect-multiplier 9
[SwitchA-LoopBack1] quit
[SwitchA] ip route-static 120.1.1.0 24 2.2.2.9 bfd control-packet bfd-source 1.1.1.9
[SwitchA] ip route-static 120.1.1.0 24 vlan-interface 11 10.1.1.100 preference 65
[SwitchA] quit
```

Configure static routes on Switch B and enable BFD control packet mode for the static route through Switch D.

```
<SwitchB> system-view
[SwitchB] interface loopback 1
[SwitchB-LoopBack1] bfd min-transmit-interval 500
[SwitchB-LoopBack1] bfd min-receive-interval 500
[SwitchB-LoopBack1] bfd detect-multiplier 9
[SwitchB-LoopBack1] quit
[SwitchB] ip route-static 121.1.1.0 24 1.1.1.9 bfd control-packet bfd-source 2.2.2.9
[SwitchB] ip route-static 121.1.1.0 24 vlan-interface 13 13.1.1.2 preference 65
[SwitchB] quit
```

Configure static routes on Switch C.

```
<SwitchC> system-view
[SwitchC] ip route-static 120.1.1.0 24 vlan-interface 13 13.1.1.1
```

```
[SwitchC] ip route-static 121.1.1.0 24 vlan-interface 11 10.1.1.102
```

Configure static routes on Switch D.

```
<SwitchD> system-view
```

```
[SwitchD] ip route-static 120.1.1.0 24 vlan-interface 12 11.1.1.1
```

```
[SwitchD] ip route-static 121.1.1.0 24 vlan-interface 10 12.1.1.1
```

3. Verify the configuration:

The following operations are performed on Switch A. The operations on Switch B are similar and not shown.

Display BFD sessions on Switch A.

```
<SwitchA> display bfd session
```

```
Total Session Num: 1          Init Mode: Active
```

```
Session Working Under Ctrl Mode:
```

LD/RD	SourceAddr	DestAddr	State	Holdtime	Interface
4/7	1.1.1.9	2.2.2.9	Up	2000ms	Loop1

Display the static route information on Switch A.

```
<SwitchA> display ip routing-table protocol static
```

```
Public Routing Table : Static
```

```
Summary Count : 2
```

```
Static Routing table Status : < Active>
```

```
Summary Count : 1
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
120.1.1.0/24	Static	60	0	2.2.2.9	Vlan10

```
Static Routing table Status : <Inactive>
```

```
Summary Count : 1
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
120.1.1.0/24	Static	65	0	10.1.1.100	Vlan11

Enable BFD debugging on Switch A. When the link between Switch A and Switch D fails, Switch A can detect the failure.

```
<SwitchA> debugging bfd event
```

```
<SwitchA> debugging bfd scm
```

```
<SwitchA> terminal debugging
```

```
%Oct 10 10:18:18:672 2010 SwitchA BFD/4/LOG:Sess[1.1.1.9/2.2.2.9, Loop1,Ctrl], Sta:  
UP->DOWN, Diag: 1
```

```
* Oct 10 10:18:18:672 2010 SwitchA BFD/7/EVENT:Send sess-down Msg,  
[Src:1.1.1.9,Dst:2.2.2.9, Loop1,Ctrl], instance:0, protocol:STATIC
```

Display the static route information on Switch A again. Switch A communicates with Switch B over the static route passing Switch C now.

```
<SwitchA> display ip routing-table protocol static
```

```
Public Routing Table : Static
```

```
Summary Count : 2
```

```
Static Routing table Status : < Active>
```

```
Summary Count : 1
```

```
Destination/Mask    Proto  Pre  Cost           NextHop           Interface
120.1.1.0/24       Static 65   0             10.1.1.100        Vlan11
Static Routing table Status : < Inactive>
Summary Count : 1
Destination/Mask    Proto  Pre  Cost           NextHop           Interface
120.1.1.0/24       Static 60   0             2.2.2.9
```

Configuring RIP

Hardware compatibility

The HP 5500 SI Switch Series does not support VPN and BFD related parameters or FRR.

Overview

RIP is a simple interior gateway protocol (IGP), mainly used in small-sized networks, such as academic networks and simple LANs. It is not applicable to complex networks.

RIP has been widely used because it is easy to implement, configure, and maintain.

The term "router" in this chapter refers to both routers and Layer 3 switches.

Understanding RIP

RIP is a distance vector (D-V) routing protocol, using UDP packets for exchanging information through port 520.

RIP uses a hop count to measure the distance to a destination. The hop count from a router to a directly connected network is 0. The hop count from a router to a directly connected router is 1. To limit convergence time, the RIP metric value is in the range of 0 to 15. A metric value of 16 (or greater) is considered infinite, which means the destination network is unreachable. Because of this, RIP is not suitable for large-sized networks.

RIP prevents routing loops by implementing the split horizon and poison reverse functions.

RIP routing table

A RIP router has a routing table containing routing entries of all reachable destinations, and each routing entry contains the following elements:

- **Destination address**—IP address of a host or a network.
- **Next hop**—IP address of the adjacent router's interface to reach the destination.
- **Egress interface**—Packet outgoing interface.
- **Metric**—Cost from the local router to the destination.
- **Route time**—Time elapsed since the last routing entry update. The time is reset to 0 every time the routing entry is updated.
- **Route tag**—Identifies a route. Route tags are used in a routing policy to flexibly control routes. For more information about routing policy, see "[Configuring routing policies](#)."

RIP timers

RIP employs the following timers:

- **Update timer**—Defines the interval between routing updates.
- **Timeout timer**—Defines the route aging time. If no update for a route is received within the aging time, the metric of the route is set to 16 in the routing table.

- **Suppress timer**—Defines how long a RIP route stays in suppressed state. When the metric of a route is 16, the route enters the suppressed state. In suppressed state, only routes coming from the same neighbor and whose metric is less than 16 will be received by the router to replace unreachable routes.
- **Garbage-collect timer**—Defines the interval from when the metric of a route becomes 16 to when it is deleted from the routing table. During the garbage-collect timer length, RIP advertises the route with the routing metric set to 16. If no update is announced for that route after the garbage-collect timer expires, then the route is deleted from the routing table.

Routing loops prevention

RIP is a distance vector routing protocol. Because a RIP router advertises its own routing table to neighbors, routing loops may occur.

RIP uses the following mechanisms to prevent routing loops:

- **Counting to infinity**—The metric value of 16 is defined as unreachable. When a routing loop occurs, the metric value of the route will increment to 16.
- **Split horizon**—A router does not send routing information back to the neighboring router the information came from, preventing routing loops and saving bandwidth.
- **Poison reverse**—A router sets the metric of routes received from a neighbor to 16 and sends back these routes to the neighbor to help delete such information from the neighbor's routing table.
- **Triggered updates**—A router advertises updates once the metric of a route is changed instead of after the update period expires to speed up network convergence.

RIP operation

The following procedure describes how RIP works:

1. After RIP is enabled, the router sends request messages to neighboring routers. Neighboring routers return response messages, including information about their routing tables.
2. After receiving this information, the router updates its local routing table, and sends triggered update messages to its neighbors. All routers on the network do this to keep the latest routing information.
3. By default, a RIP router sends its routing table to neighbors every 30 seconds.
4. RIP ages out routes by adopting an aging mechanism to keep only valid routes.

RIP versions

RIP has RIPv1 and RIPv2.

RIPv1, a classful routing protocol, supports message advertisement through broadcast only. RIPv1 protocol messages do not carry mask information, so it can only recognize routing information of natural networks such as Class A, B, and C. For this reason, RIPv1 does not support discontinuous subnets.

RIPv2 is a classless routing protocol and has the following advantages over RIPv1.

- Supports route tags, which are used in routing policies to flexibly control routes.
- Supports masks, route summarization, and Classless Inter-Domain Routing (CIDR).
- Supports designated next hops to select the best ones on broadcast networks.
- Supports multicasting routing updates to reduce resource consumption. Only RIPv2 routers can receive these update messages.

- Supports simple authentication and MD5 authentication to enhance security.

NOTE:

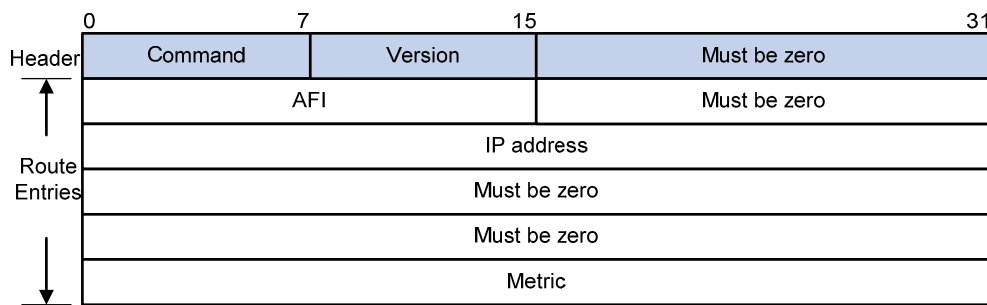
RIPv2 has two types of message transmission: broadcast and multicast. Multicast is the default type using 224.0.0.9 as the multicast address. The interface operating in the RIPv2 broadcast mode can also receive RIPv1 messages.

RIP message format

A RIP message consists of a header and up to 25 route entries. (A RIPv2 authentication message uses the first route entry as the authentication entry, leaving 24 available.)

RIPv1 message format

Figure 6 RIPv1 message format

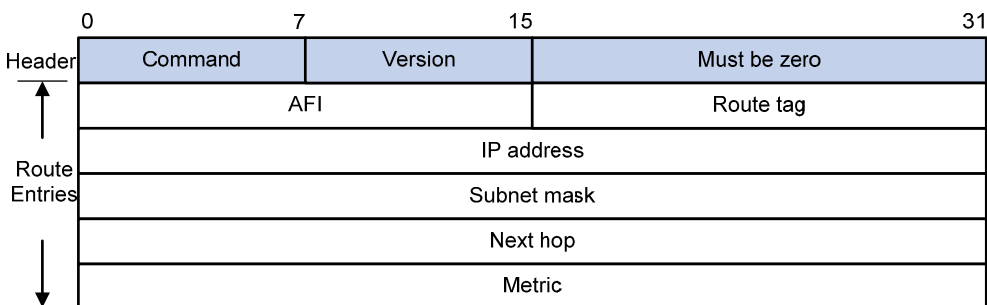


- **Command**—Type of message. 1 indicates request, which is used to request all or part of the routing information from the neighbor; 2 indicates response, which contains all or part of the routing information. A response message consists of up to 25 route entries.
- **Version**—Version of RIP, 0x01 for RIPv1.
- **Must be zero**—This field must be zero
- **AFI**—Address Family Identifier, 2 for IP.
- **IP address**—Destination IP address of the route that can be a natural network, subnet, or a host address.
- **Metric**—Cost of the route.

RIPv2 message format

The format of RIPv2 message is similar to RIPv1.

Figure 7 RIPv2 Message Format



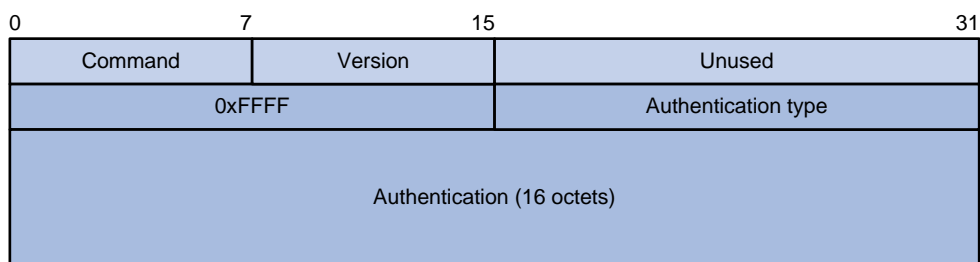
The differences between RIPv1 and RIPv2 messages are stated as follows:

- **Version**—Version of RIP. For RIPv2 the value is 0x02.
- Route tag.
- **IP address**—Destination IP address. It can be a natural network address, subnet address, or host address.
- **Subnet mask**—Mask of the destination address. Unlike RIPv1, RIPv2 can carry subnet information.
- **Next hop**—If set to 0.0.0.0, it indicates that the originator of the route is the best next hop. Otherwise, it indicates a next hop better than the originator of the route.

RIPv2 authentication message format

RIPv2 sets the AFI field of the first route entry to 0xFFFF to identify authentication information.

Figure 8 RIPv2 authentication message



- **Authentication Type**—A value of 2 represents simple authentication. A value of 3 represents MD5. RFC 1723 only defines simple authentication. For more information about MD5 authentication, see RFC 2453, *RIP Version 2*.
- **Authentication**—Authentication data, including password information when simple authentication is adopted or including key ID, MD5 authentication data length and sequence number when MD5 authentication is adopted.

NOTE:

Even though with RIPv1, you can configure the authentication mode in interface view, the configuration will not take effect because RIPv1 does not support authentication.

Supported RIP features

The current implementation supports the following RIP features:

- RIPv1 and RIPv2
- RIP support for multi-VPN-instance
- RIP FRR
- BFD

RIP periodically sends route update requests to neighbors. If no route update response for a route is received within the specified interval, RIP considers the route unreachable. This mechanism cannot detect link faults quickly. After BFD is configured for RIP, and BFD detects a broken link, RIP can quickly age out the unreachable route, thus avoiding interference to other services.

Protocols and standards

- RFC 1058, *Routing Information Protocol*
- RFC 1723, *RIP Version 2 - Carrying Additional Information*
- RFC 1721, *RIP Version 2 Protocol Analysis*
- RFC 1722, *RIP Version 2 Protocol Applicability Statement*
- RFC 1724, *RIP Version 2 MIB Extension*
- RFC 2082, *RIPv2 MD5 Authentication*
- RFC 2453, *RIP Version 2*

RIP configuration task list

Task	Remarks	
Configuring RIP basic functions	Required	
Configuring RIP route control	Configuring an additional routing metric	Optional
	Configuring RIPv2 route summarization	Optional
	Disabling host route reception	Optional
	Advertising a default route	Optional
	Configuring inbound or outbound route filtering	Optional
	Configuring a priority for RIP	Optional
	Configuring RIP route redistribution	Optional
Tuning and optimizing RIP networks	Configuring RIP timers	Optional
	Configuring split horizon and poison reverse	Optional
	Configuring the maximum number of ECMP routes	Optional
	Enabling zero field check on incoming RIPv1 messages	Optional
	Enabling source IP address check on incoming RIP updates	Optional
	Configuring RIPv2 message authentication	Optional
	Specifying a RIP neighbor	Optional
	Configuring RIP-to-MIB binding	Optional
Configuring the RIP packet sending rate	Optional	
Configuring RIP FRR	Optional	
Configuring BFD for RIP	Single-hop detection in BFD echo packet mode	Optional
	Bidirectional detection in BFD control packet mode	Optional

Configuring RIP basic functions

Configuration prerequisites

Before you configure RIP basic functions, complete the following tasks:

- Configure the link layer protocol.
- Configure IP addresses for interfaces, and make sure that all neighboring routers can reach each other.

Enabling RIP and a RIP interface

Follow these guidelines when you enable RIP:

- RIP configurations made in interface view before enabling RIP take effect after RIP is enabled.
- RIP runs only on the interfaces residing on the specified networks. Specify the network after enabling RIP to validate RIP on a specific interface.
- You can enable RIP on all interfaces using the command **network** 0.0.0.0.
- If a physical interface is attached to multiple networks, you cannot advertise these networks in different RIP processes.

To enable RIP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable a RIP process and enter RIP view.	rip [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	Disabled by default.
3. Enable RIP on the interface attached to the specified network.	network <i>network-address</i>	Disabled by default.

Configuring the interface behavior

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIP view.	rip [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Disable an or all interfaces from sending routing updates (the interfaces can still receive updates).	silent-interface { <i>interface-type</i> <i>interface-number</i> all }	Optional. By default, all interfaces can send routing updates.
4. Return to system view.	quit	N/A
5. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
6. Enable the interface to receive RIP messages.	rip input	Optional. Enabled by default.

Step	Command	Remarks
7. Enable the interface to send RIP messages.	rip output	Optional. Enabled by default.

Configuring a RIP version

You can configure a RIP version in RIP view or interface view under the following conditions.

- If neither global nor interface RIP version is configured, the interface sends RIPv1 broadcasts and can receive the following packets: RIPv1 broadcast, RIPv1 unicast, RIPv2 broadcast, RIPv2 multicast, and RIPv2 unicast.
- If an interface has no RIP version configured, it uses the global RIP version; otherwise it uses the RIP version configured on it.
- With RIPv1 configured, an interface sends RIPv1 broadcasts, and can receive RIPv1 broadcasts and RIPv1 unicasts.
- With RIPv2 configured, a multicast interface sends RIPv2 multicasts and can receive RIPv2 unicasts, broadcasts, and multicasts.
- With RIPv2 configured, a broadcast interface sends RIPv2 broadcasts and can receive RIPv1 unicasts and broadcasts, and RIPv2 broadcasts, multicasts, and unicasts.

To configure a RIP version:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIP view.	rip [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Specify a global RIP version.	version { 1 2 }	Optional. By default, if an interface has a RIP version specified, the version takes precedence over the global one. If no RIP version is specified for an interface, the interface can send RIPv1 broadcasts, and receive RIPv1 broadcasts and unicasts, and RIPv2 broadcasts, multicasts, and unicasts.
4. Return to system view.	quit	N/A
5. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
6. Specify a RIP version for the interface.	rip version { 1 2 [broadcast multicast] }	Optional. By default, if an interface has no RIP version specified, the global version takes effect. If no global RIP version is specified, the interface can send RIPv1 broadcasts, and receive RIPv1 broadcasts and unicasts, and RIPv2 broadcasts, multicasts, and unicasts.

Setting the DSCP value for RIP packets

An IPv4 packet header contains an 8-bit Type of Service (ToS) field. As defined in RFC 2474, the first six bits set the Differentiated Services Code Point (DSCP) value, and the last two bits are reserved. Network devices use the DSCP value as a reference to determine the packet priority for transmission.

You can set the DSCP value for RIP packets.

To set the DSCP value for RIP packets:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a RIP process and enter RIP view.	rip [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	By default, no RIP process is created.
3. Set the DSCP value for RIP packets.	dscp <i>dscp-value</i>	Optional. By default, the DSCP value in RIP packets is 48.

Configuring RIP route control

In complex networks, you must configure advanced RIP features.

Before you configure RIP routing feature, complete the following tasks:

- Configure IP addresses for interfaces, and make sure that all neighboring routers can reach each other.
- Configure RIP basic functions.

Configuring an additional routing metric

An additional routing metric (hop count) can be added to the metric of an inbound or outbound RIP route.

The outbound additional metric is added to the metric of a sent route, and the route's metric in the routing table is not changed.

The inbound additional metric is added to the metric of a received route before the route is added into the routing table, and the route's metric is changed. If the sum of the additional metric and the original metric is greater than 16, the metric of the route will be 16.

To configure additional routing metrics:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Define an inbound additional routing metric.	rip metricin [route-policy <i>route-policy-name</i>] <i>value</i>	Optional. 0 by default.
4. Define an outbound additional routing metric.	rip metricout [route-policy <i>route-policy-name</i>] <i>value</i>	Optional. 1 by default.

Configuring RIPv2 route summarization

Route summarization means that subnets in a natural network are summarized into a natural network that is sent to other networks. This feature can reduce the size of routing tables.

After route summarization, the smallest metric among all the summarized routes is used as the metric of the summary route.

Enabling RIPv2 route automatic summarization

You can disable RIPv2 route automatic summarization if you want to advertise all subnet routes.

To enable RIPv2 route automatic summarization:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIP view.	rip [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Enable RIPv2 automatic route summarization.	summary	Optional. Enabled by default. If the subnet routes in the routing table are not consecutive, disable automatic route summarization to avoid black hole routing.

Advertising a summary route

Disable RIPv2 route automatic summarization before advertising a summary route on an interface.

To configure RIPv2 to advertise a summary route on the specified interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIP view.	rip [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Disable RIPv2 automatic route summarization.	undo summary	Enabled by default.
4. Return to system view.	quit	N/A
5. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
6. Advertise a summary route.	rip summary-address <i>ip-address</i> { <i>mask</i> <i>mask-length</i> }	N/A

Disabling host route reception

Perform this task to disable RIPv2 from receiving host routes from the same network and save network resources. This feature does not apply to RIPv1.

To disable RIP from receiving host routes:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIP view.	rip [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Disable RIP from receiving host routes.	undo host-route	Enabled by default.

Advertising a default route

Under the following conditions, you can configure RIP to advertise a default route with a specified metric to RIP neighbors:

- In RIP view, you can configure all the interfaces of the RIP process to advertise a default route; in interface view, you can configure a RIP interface of the RIP process to advertise a default route. The interface view takes precedence over on the interface.
- If a RIP process is enabled to advertise a default route, to disable an interface of the RIP process from default route advertisement, you can use the **rip default-route no-originate** command on the interface.

To configure RIP to advertise a default route:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIP view.	rip [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Enable RIP to advertise a default route.	default-route { only originate } [cost <i>cost</i>]	Optional. Not enabled by default.
4. Return to system view.	quit	N/A
5. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
6. Configure the RIP interface to advertise a default route.	rip default-route { { only originate } [cost <i>cost</i>] no-originate }	Optional. By default, a RIP interface can advertise a default route if the RIP process is configured with default route advertisement.

NOTE:

The router enabled to advertise a default route does not receive default routes from RIP neighbors.

Configuring inbound or outbound route filtering

The device supports route filtering. You can filter routes by configuring the inbound and outbound route filtering policies by referencing an ACL or IP prefix list. You can also configure the router to receive only routes from a specified neighbor.

To configure route filtering:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIP view.	rip [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure the filtering of incoming routes.	filter-policy { <i>acl-number</i> gateway <i>ip-prefix-name</i> ip-prefix <i>ip-prefix-name</i> [gateway <i>ip-prefix-name</i>] } import [<i>interface-type interface-number</i>]	Not configured by default. The filter-policy import command filters incoming routes. Routes not passing the filtering will be neither installed into the routing table nor advertised to neighbors.
4. Configure the filtering of outgoing routes.	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } export [<i>protocol</i> [<i>process-id</i>] <i>interface-type interface-number</i>]	Not configured by default. The filter-policy export command filters outgoing routes, including routes redistributed with the import-route command.

Configuring a priority for RIP

Multiple IGP protocols can run in a router. If you want RIP routes to have a higher priority than those learned by other routing protocols, then assign RIP a smaller priority value to influence optimal route selection.

To configure a priority for RIP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIP view.	rip [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure a priority for RIP.	preference [route-policy <i>route-policy-name</i>] <i>value</i>	Optional. 100 by default.

Configuring RIP route redistribution

If a router runs RIP and other routing protocols, you can configure RIP to redistribute OSPF, IS-IS, BGP, static, or direct routes.

Only active routes can be redistributed. To display route state information, use the **display ip routing-table protocol** command.

To configure RIP route redistribution:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIP view.	rip [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A

Step	Command	Remarks
3. Configure a default metric for redistributed routes.	default cost <i>value</i>	Optional. The default metric of a redistributed route is 0.
4. Redistribute routes from another protocol.	import-route <i>protocol</i> [<i>process-id</i> all-processes allow-ibgp] [cost <i>cost</i> route-policy <i>route-policy-name</i> tag <i>tag</i>] *	By default, no redistribution is configured.

Tuning and optimizing RIP networks

Before you tune and optimize RIP networks, complete the following tasks:

- Configure IP addresses for interfaces, and make sure that all neighboring nodes can reach each other.
- Configure RIP basic functions.

Configuring RIP timers

You can change the RIP network convergence speed by adjusting RIP timers. Based on network performance, make RIP timers of RIP routers identical to each other to avoid unnecessary traffic or route oscillation.

To configure RIP timers:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIP view.	rip [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure values for RIP timers.	timers { garbage-collect <i>garbage-collect-value</i> suppress <i>suppress-value</i> timeout <i>timeout-value</i> update <i>update-value</i> } *	Optional. The default update timer, timeout timer, suppress timer, and garbage-collect timer are 30s, 180s, 120s and 120s respectively.

Configuring split horizon and poison reverse

The split horizon and poison reverse functions can avoid routing loops. If both split horizon and poison reverse are configured, only the poison reverse function takes effect.

Enabling split horizon

The split horizon function disables an interface from sending routes received from the interface to prevent routing loops between adjacent routers.

To enable split horizon:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable split horizon.	rip split-horizon	Optional. Enabled by default.

Enabling poison reverse

The poison reverse function allows an interface to advertise the routes received from it, but the metric of these routes is set to 16, making them unreachable. This can avoid routing loops between neighbors.

To enable poison reverse:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable poison reverse.	rip poison-reverse	Disabled by default.

Configuring the maximum number of ECMP routes

Perform this task to implement load sharing over ECMP routes.

To configure the maximum number of ECMP routes:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIP view.	rip [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure the maximum number of ECMP routes.	maximum load-balancing <i>number</i>	Optional. 8 by default.

Enabling zero field check on incoming RIPv1 messages

Some fields in the RIPv1 message must be zero. These fields are called "zero fields." You can enable zero field check on received RIPv1 messages. If such a field contains a non-zero value, the RIPv1 message will not be processed. If you are certain that all messages are trustworthy, then disable zero field check to save CPU resources.

This feature does not apply to RIPv2 packets, because they have no zero fields.

To enable zero field check on incoming RIPv1 messages:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter RIP view.	rip [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Enable zero field check on received RIPv1 messages.	checkzero	Optional. Enabled by default.

Enabling source IP address check on incoming RIP updates

You can enable source IP address check on incoming RIP updates.

For a message received on an Ethernet interface, RIP compares the source IP address of the message with the IP address of the interface. If they are not in the same network segment, RIP discards the message.

ⓘ IMPORTANT:

Disable the source IP address check feature if the RIP neighbor is not directly connected.

To enable source IP address check on incoming RIP updates:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIP view.	rip [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Enable source IP address check on incoming RIP messages.	validate-source-address	Optional. Enabled by default.

Configuring RIPv2 message authentication

In a network requiring high security, configure this task to implement RIPv2 message validity check and authentication. This feature does not apply to RIPv1 because RIPv1 does not support authentication. Although you can specify an authentication mode for RIPv1 in interface view, the configuration does not take effect.

RIPv2 supports simple authentication and MD5 authentication.

To configure RIPv2 message authentication:

Step	Command
1. Enter system view.	system-view
2. Enter interface view.	interface <i>interface-type interface-number</i>
3. Configure authentication.	RIPv2 rip authentication-mode { md5 { rfc2082 [cipher] <i>key-string key-id</i> rfc2453 [cipher] <i>key-string</i> } simple [cipher] <i>password</i> }

Specifying a RIP neighbor

Usually, RIP sends messages to broadcast or multicast addresses. On non-broadcast or multicast links, you must manually specify RIP neighbors.

Follow these guidelines when you specify a RIP neighbor:

- Do not use the **peer ip-address** command when the neighbor is directly connected because the neighbor may receive both the unicast and multicast (or broadcast) of the same routing information.
- If a specified neighbor is not directly connected, then disable the source address check on incoming updates.

To specify a RIP neighbor:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIP view.	rip [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Specify a RIP neighbor.	peer ip-address	N/A
4. Disable source address check on incoming RIP updates.	undo validate-source-address	Not disabled by default.

Configuring RIP-to-MIB binding

This task allows you to enable a specific RIP process to receive SNMP requests.

To bind RIP to MIB:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Bind RIP to MIB.	rip mib-binding <i>process-id</i>	Optional. By default, MIB is bound to RIP process 1.

Configuring the RIP packet sending rate

RIP periodically sends routing information in RIP packets to RIP neighbors.

Sending large numbers of RIP packets at the same time may affect device performance and consume large network bandwidth. To solve this problem, specify the maximum number of RIP packets that can be sent at the specified interval.

To configure the RIP packet sending rate:

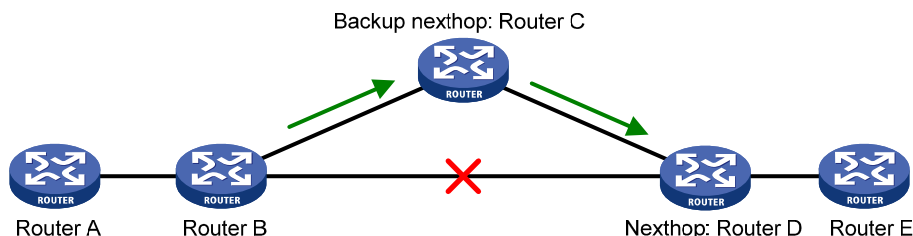
Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable a RIP process and enter RIP view.	rip [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure the maximum number of RIP packets that can be sent at the specified interval.	output-delay <i>time count</i> <i>count</i>	Optional. By default, an interface sends up to three RIP packets every 20 milliseconds.

Configuring RIP FRR

When a link in a RIP network fails, the traffic is interrupted until RIP completes routing convergence based on the new network topology.

You can enable RIP fast reroute (FRR) to reduce traffic recovery time.

Figure 9 Network diagram for RIP FRR



In [Figure 9](#), after you enable FRR on Router B, RIP designates a backup next hop using a routing policy when a network failure is detected. Packets are directed to the backup next hop to reduce traffic recovery time. At the same time, RIP calculates the shortest path based on the new network topology, and forwards packets over the path after network convergence.

Configuration prerequisites

You need to specify a next hop by using the **apply fast-reroute backup-interface** command in a routing policy and reference the routing policy with RIP FRR. For more information about routing policy configuration, see "[Configuring routing policies.](#)"

Configuration guidelines

- RIP FRR is only effective for non-recursive RIP routes (that are learned from directly connected neighbors).
- Do not use RIP FRR and BFD (for RIP) at the same time; otherwise, RIP FRR may fail to take effect.

Configuration procedure

To configure RIP FRR:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the source address of echo packets.	bfd echo-source-ip ip-address	Not configured by default.
3. Enter RIP view.	rip [process-id] [vpn-instance vpn-instance-name]	N/A
4. Enable RIP FRR and reference a routing policy to designate a backup next hop.	fast-reroute route-policy route-policy-name	Disabled by default.

Configuring BFD for RIP

BFD for RIP provides the following link detection modes:

- Single-hop detection in BFD echo packet mode for a directly connected neighbor. In this mode, a BFD session is established only when the neighbor has route information to send.
- Bidirectional detection in BFD control packet mode for an indirectly connected neighbor. In this mode, a BFD session is established only when both ends have routes to send and BFD is enabled on the receiving interface.

For more information about BFD, see *High Availability Configuration Guide*.

Single-hop detection in BFD echo packet mode

To configure BFD for RIP (single-hop detection in BFD echo packet mode):

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the source IP address of BFD echo packets.	bfd echo-source-ip <i>ip-address</i>	By default, no source IP address is configured for BFD echo packets.
3. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Enable BFD on the RIP interface.	rip bfd enable	Disabled by default.

Bidirectional detection in BFD control packet mode

This feature only works for RIP neighbors that are directly connected (one hop away from each other).

To configure BFD for RIP (bidirectional detection in BFD control packet mode):

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a RIP process and enter RIP view.	rip [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	By default, RIP is disabled.
3. Specify a RIP neighbor.	peer <i>ip-address</i>	By default, RIP does not unicast updates to any peer.
4. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
5. Enable BFD on the RIP interface.	rip bfd enable	Disabled by default.

NOTE:

Because the **undo peer** command does not remove the neighbor relationship at once, executing the command cannot bring down the BFD session at once.

Displaying and maintaining RIP

Task	Command	Remarks
Display RIP current status and configuration information.	display rip [<i>process-id</i> vpn-instance <i>vpn-instance-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display all active routes in RIP database.	display rip <i>process-id</i> database [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display RIP interface information.	display rip <i>process-id</i> interface [<i>interface-type</i> <i>interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display routing information about a specified RIP process.	display rip <i>process-id</i> route [<i>ip-address</i> { <i>mask</i> <i>mask-length</i> } peer <i>ip-address</i> statistics] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Reset a RIP process.	reset rip <i>process-id</i> process	Available in user view
Clear the statistics of a RIP process.	reset rip <i>process-id</i> statistics	Available in user view

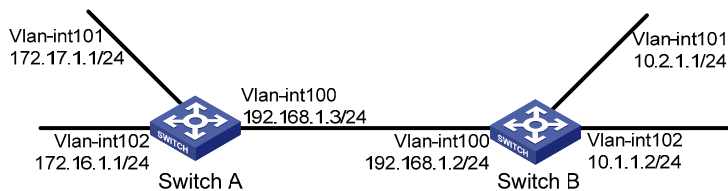
RIP configuration examples

Configuring RIP version

Network requirements

As shown in [Figure 10](#), enable RIPv2 on all interfaces on Switch A and Switch B.

Figure 10 Network diagram



Configuration procedure

1. Configure an IP address for each interface. (Details not shown.)
2. Configure basic RIP functions:

Configure Switch A.

```
[SwitchA] rip
[SwitchA-rip-1] network 192.168.1.0
[SwitchA-rip-1] network 172.16.0.0
[SwitchA-rip-1] network 172.17.0.0
```

Configure Switch B.

```
[SwitchB] rip
[SwitchB-rip-1] network 192.168.1.0
```

```
[SwitchB-rip-1] network 10.0.0.0
# Display the RIP routing table of Switch A.
[SwitchA] display rip 1 route
Route Flags: R - RIP, T - TRIP
              P - Permanent, A - Aging, S - Suppressed, G - Garbage-collect
-----
Peer 192.168.1.2 on Vlan-interface100
  Destination/Mask      Nexthop      Cost      Tag      Flags      Sec
  10.0.0.0/8           192.168.1.2    1         0        RA         11
```

The output shows that RIPv1 uses a natural mask.

3. Configure RIP version:

Configure RIPv2 on Switch A.

```
[SwitchA] rip
[SwitchA-rip-1] version 2
[SwitchA-rip-1] undo summary
```

Configure RIPv2 on Switch B.

```
[SwitchB] rip
[SwitchB-rip-1] version 2
[SwitchB-rip-1] undo summary
```

Display the RIP routing table on Switch A.

```
[SwitchA] display rip 1 route
Route Flags: R - RIP, T - TRIP
              P - Permanent, A - Aging, S - Suppressed, G - Garbage-collect
-----
Peer 192.168.1.2 on Vlan-interface100
  Destination/Mask      Nexthop      Cost      Tag      Flags      Sec
  10.0.0.0/8           192.168.1.2    1         0        RA         50
  10.2.1.0/24          192.168.1.2    1         0        RA         16
  10.1.1.0/24          192.168.1.2    1         0        RA         16
```

The output shows that RIPv2 uses classless subnet mask.

NOTE:

RIPv1 routing information has a long aging time, so it will exist until it ages out after RIPv2 is configured.

Configuring RIP route redistribution

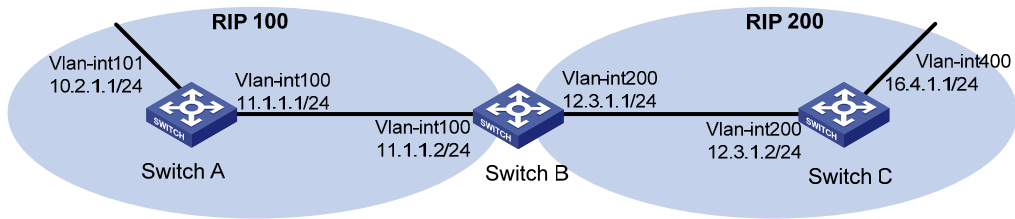
Network requirements

In the following figure, two RIP processes are running on Switch B, which communicates with Switch A through RIP 100 and with Switch C through RIP 200.

Configure route redistribution on Switch B to make RIP 200 redistribute direct routes and routes from RIP 100. Switch C can then learn routes destined for 10.2.1.0/24 and 11.1.1.0/24, and Switch A cannot learn routes destined for 12.3.1.0/24 and 16.4.1.0/24.

Configure a filtering policy on Switch B to filter out the route 10.2.1.1/24 from RIP 100, making the route not advertised to Switch C.

Figure 11 Network diagram



Configuration procedure

1. Configure an IP address for each interface. (Details not shown.)
2. Configure basic RIP functions:

Enable RIP 100 and specify RIP version 2 on Switch A.

```
<SwitchA> system-view
[SwitchA] rip 100
[SwitchA-rip-100] network 10.0.0.0
[SwitchA-rip-100] network 11.0.0.0
[SwitchA-rip-100] version 2
[SwitchA-rip-100] undo summary
[SwitchA-rip-100] quit
```

Enable RIP 100 and RIP 200 and specify RIP version 2 on Switch B.

```
<SwitchB> system-view
[SwitchB] rip 100
[SwitchB-rip-100] network 11.0.0.0
[SwitchB-rip-100] version 2
[SwitchB-rip-100] undo summary
[SwitchB-rip-100] quit
[SwitchB] rip 200
[SwitchB-rip-200] network 12.0.0.0
[SwitchB-rip-200] version 2
[SwitchB-rip-200] undo summary
[SwitchB-rip-200] quit
```

Enable RIP 200 and specify RIP version 2 on Switch C.

```
<SwitchC> system-view
[SwitchC] rip 200
[SwitchC-rip-200] network 12.0.0.0
[SwitchC-rip-200] network 16.0.0.0
[SwitchC-rip-200] version 2
[SwitchC-rip-200] undo summary
```

Display the routing table of Switch C.

```
[SwitchC] display ip routing-table
```

Routing Tables: Public

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
12.3.1.0/24	Direct	0	0	12.3.1.2	Vlan200
12.3.1.2/32	Direct	0	0	127.0.0.1	InLoop0
16.4.1.0/24	Direct	0	0	16.4.1.1	Vlan400

```

16.4.1.1/32          Direct 0    0          127.0.0.1      InLoop0
127.0.0.0/8         Direct 0    0          127.0.0.1      InLoop0
127.0.0.1/32        Direct 0    0          127.0.0.1      InLoop0

```

3. Configure route redistribution:

On Switch B, configure RIP 200 to redistribute direct routes and routes from RIP 100.

```

[SwitchB] rip 200
[SwitchB-rip-200] import-route rip 100
[SwitchB-rip-200] import-route direct
[SwitchB-rip-200] quit

```

Display the routing table of Switch C.

```

[SwitchC] display ip routing-table
Routing Tables: Public
          Destinations : 8          Routes : 8
Destination/Mask    Proto  Pre  Cost           NextHop         Interface
10.2.1.0/24         RIP    100  1             12.3.1.1        Vlan200
11.1.1.0/24         RIP    100  1             12.3.1.1        Vlan200
12.3.1.0/24         Direct 0    0             12.3.1.2        Vlan200
12.3.1.2/32         Direct 0    0             127.0.0.1       InLoop0
16.4.1.0/24         Direct 0    0             16.4.1.1        Vlan400
16.4.1.1/32         Direct 0    0             127.0.0.1       InLoop0
127.0.0.0/8         Direct 0    0             127.0.0.1       InLoop0
127.0.0.1/32        Direct 0    0             127.0.0.1       InLoop0

```

4. Configure an filtering policy to filter redistributed routes:

Define ACL 2000 and reference it to a filtering policy to filter routes redistributed from RIP 100 on Switch B, making the route not advertised to Switch C.

```

[SwitchB] acl number 2000
[SwitchB-acl-basic-2000] rule deny source 10.2.1.1 0.0.0.255
[SwitchB-acl-basic-2000] rule permit
[SwitchB-acl-basic-2000] quit
[SwitchB] rip 200
[SwitchB-rip-200] filter-policy 2000 export rip 100

```

Display the routing table of Switch C.

```

[SwitchC] display ip routing-table
Routing Tables: Public
          Destinations : 7          Routes : 7
Destination/Mask    Proto  Pre  Cost           NextHop         Interface
11.1.1.0/24         RIP    100  1             12.3.1.1        Vlan200
12.3.1.0/24         Direct 0    0             12.3.1.2        Vlan200
12.3.1.2/32         Direct 0    0             127.0.0.1       InLoop0
16.4.1.0/24         Direct 0    0             16.4.1.1        Vlan400
16.4.1.1/32         Direct 0    0             127.0.0.1       InLoop0
127.0.0.0/8         Direct 0    0             127.0.0.1       InLoop0
127.0.0.1/32        Direct 0    0             127.0.0.1       InLoop0

```

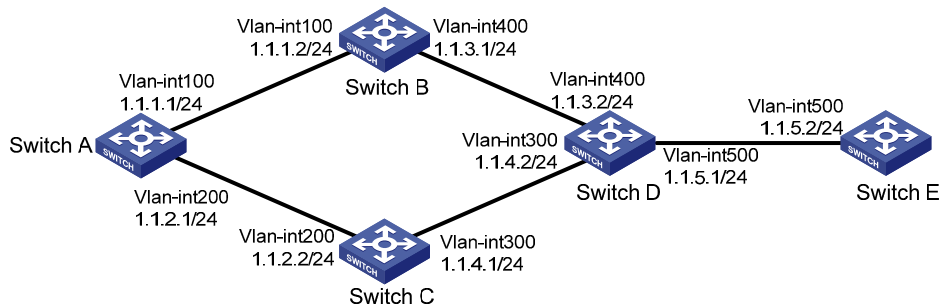
Configuring an additional metric for a RIP interface

Network requirements

In the following figure, RIP is enabled on all the interfaces of Switch A, Switch B, Switch C, Switch D, and Switch E. The switches are interconnected through RIPv2.

Switch A has two links to Switch D. The link from Switch B to Switch D is more stable than that from Switch C to Switch D. Configure an additional metric for RIP routes received through VLAN-interface 200 on Switch A so that Switch A prefers the 1.1.5.0/24 network learned from Switch B.

Figure 12 Network diagram



Configuration procedure

1. Configure IP addresses for the interfaces. (Details not shown.)
2. Configure RIP basic functions:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] rip 1
[SwitchA-rip-1] network 1.0.0.0
[SwitchA-rip-1] version 2
[SwitchA-rip-1] undo summary
[SwitchA-rip-1] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] rip 1
[SwitchB-rip-1] network 1.0.0.0
[SwitchB-rip-1] version 2
[SwitchB-rip-1] undo summary
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] rip 1
[SwitchC-rip-1] network 1.0.0.0
[SwitchC-rip-1] version 2
[SwitchC-rip-1] undo summary
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] rip 1
[SwitchD-rip-1] network 1.0.0.0
[SwitchD-rip-1] version 2
```

```

[SwitchD-rip-1] undo summary
# Configure Switch E.
<SwitchE> system-view
[SwitchE] rip 1
[SwitchE-rip-1] network 1.0.0.0
[SwitchE-rip-1] version 2
[SwitchE-rip-1] undo summary
# Display the IP routing table of Switch A.
[SwitchA] display rip 1 database
  1.0.0.0/8, cost 0, ClassfulSumm
    1.1.1.0/24, cost 0, nexthop 1.1.1.1, Rip-interface
    1.1.2.0/24, cost 0, nexthop 1.1.2.1, Rip-interface
    1.1.3.0/24, cost 1, nexthop 1.1.1.2
    1.1.4.0/24, cost 1, nexthop 1.1.2.2
    1.1.5.0/24, cost 2, nexthop 1.1.1.2
    1.1.5.0/24, cost 2, nexthop 1.1.2.2

```

The output shows that two RIP routes can reach network 1.1.5.0/24. Their next hops are Switch B (1.1.1.2) and Switch C (1.1.2.2) respectively, with the same cost of 2. Switch C is the next hop router to reach network 1.1.4.0/24, with a cost of 1.

3. Configure an additional metric for the RIP interface:

Configure an additional metric of 3 for VLAN-interface 200 on Switch A.

```

[SwitchA] interface vlan-interface 200
[SwitchA-Vlan-interface200] rip metricin 3
[SwitchA-Vlan-interface200] display rip 1 database
  1.0.0.0/8, cost 0, ClassfulSumm
    1.1.1.0/24, cost 0, nexthop 1.1.1.1, Rip-interface
    1.1.2.0/24, cost 0, nexthop 1.1.2.1, Rip-interface
    1.1.3.0/24, cost 1, nexthop 1.1.1.2
    1.1.4.0/24, cost 2, nexthop 1.1.1.2
    1.1.5.0/24, cost 2, nexthop 1.1.1.2

```

The output shows that only one RIP route reaches network 1.1.5.0/24, with the next hop as Switch B (1.1.1.2) and a cost of 2.

Configuring RIP to advertise a summary route

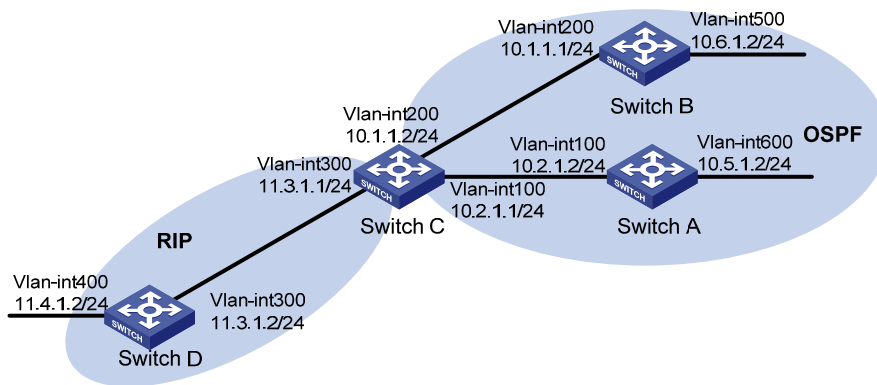
Network requirements

In the following figure, Switch A and Switch B run OSPF, Switch D runs RIP, and Switch C runs OSPF and RIP.

Configure RIP to redistribute OSPF routes on Switch C so that Switch D has routes destined for networks 10.1.1.0/24, 10.2.1.0/24, 10.5.1.0/24, and 10.6.1.0/24.

Configure route summarization on Switch C and advertise only the summary route 10.0.0.0/8, reducing the routing table size of Switch D.

Figure 13 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure OSPF basic functions:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] ospf
[SwitchA-ospf-1] area 0
[SwitchA-ospf-1-area-0.0.0.0] network 10.5.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] network 10.2.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] ospf
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] network 10.6.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] ospf
[SwitchC-ospf-1] area 0
[SwitchC-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] network 10.2.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] quit
```

3. Configure RIP basic functions:

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] rip 1
[SwitchC-rip-1] network 11.3.1.0
[SwitchC-rip-1] version 2
[SwitchC-rip-1] undo summary
```

Configure Switch D.

```
<SwitchD> system-view
```

```
[SwitchD] rip 1
[SwitchD-rip-1] network 11.0.0.0
[SwitchD-rip-1] version 2
[SwitchD-rip-1] undo summary
[SwitchD-rip-1] quit
```

Configure RIP to redistribute the routes from OSPF process 1 and direct routes on Switch C.

```
[SwitchC-rip-1] import-route direct
[SwitchC-rip-1] import-route ospf 1
```

Display the routing table information of Switch D.

```
[SwitchD] display ip routing-table
Routing Tables: Public
                Destinations : 10          Routes : 10
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
10.1.1.0/24	RIP	100	1	11.3.1.1	Vlan300
10.2.1.0/24	RIP	100	1	11.3.1.1	Vlan300
10.5.1.0/24	RIP	100	1	11.3.1.1	Vlan300
10.6.1.0/24	RIP	100	1	11.3.1.1	Vlan300
11.3.1.0/24	Direct	0	0	11.3.1.2	Vlan300
11.3.1.2/32	Direct	0	0	127.0.0.1	InLoop0
11.4.1.0/24	Direct	0	0	11.4.1.2	Vlan400
11.4.1.2/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0

4. Configure route summarization on Switch C and advertise only the summary route 10.0.0.0/8:

```
[SwitchC] interface vlan-interface 300
[SwitchC-Vlan-interface300] rip summary-address 10.0.0.0 8
```

Display the routing table information of Switch D.

```
[SwitchD] display ip routing-table
Routing Tables: Public
                Destinations : 7          Routes : 7
```

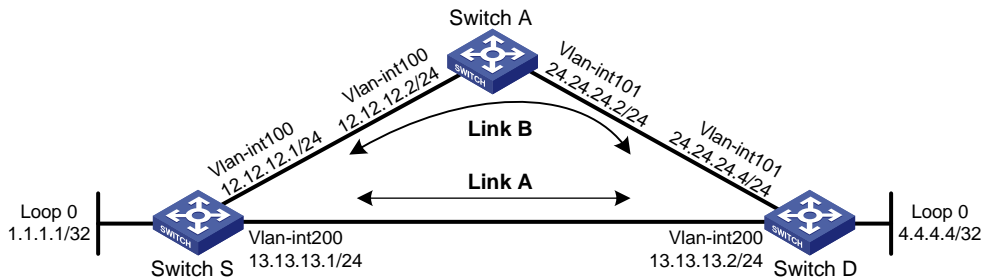
Destination/Mask	Proto	Pre	Cost	NextHop	Interface
10.0.0.0/8	RIP	100	1	11.3.1.1	Vlan300
11.3.1.0/24	Direct	0	0	11.3.1.2	Vlan300
11.3.1.2/32	Direct	0	0	127.0.0.1	InLoop0
11.4.1.0/24	Direct	0	0	11.4.1.2	Vlan400
11.4.1.2/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0

RIP FRR configuration example

Network requirements

Switch S, Switch A, and Switch D are interconnected through RIPv2, as illustrated in Figure 14. Configure RIP FRR so that when Link A between Switch S and Switch D fails, services can be switched to Link B immediately.

Figure 14 Network diagram



Configuration procedure

1. Configure IP addresses for the interfaces on each switch and configure RIPv2:
Follow Figure 14 to configure the IP address and subnet mask of each interface on the switches. (Details not shown.)
Configure RIPv2 on the switches, ensuring that Switch A, Switch D, and Switch S can communicate with each other at Layer 3. (Details not shown.)

2. Configure RIP FRR:

Configure Switch S.

```
<SwitchS> system-view
[SwitchS] bfd echo-source-ip 1.1.1.1
[SwitchS] ip ip-prefix abc index 10 permit 4.4.4.4 32
[SwitchS] route-policy frr permit node 10
[SwitchS-route-policy] if-match ip-prefix abc
[SwitchS-route-policy] apply fast-reroute backup-interface vlan-interface 100
backup-nexthop 12.12.12.2
[SwitchS-route-policy] quit
[SwitchS] rip 1
[SwitchS-rip-1] fast-reroute route-policy frr
[SwitchS-rip-1] quit
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] bfd echo-source-ip 4.4.4.4
[SwitchD] ip ip-prefix abc index 10 permit 1.1.1.1 32
[SwitchD] route-policy frr permit node 10
[SwitchD-route-policy] if-match ip-prefix abc
[SwitchD-route-policy] apply fast-reroute backup-interface vlan-interface 101
backup-nexthop 24.24.24.2
[SwitchD-route-policy] quit
[SwitchD] rip 1
[SwitchD-rip-1] fast-reroute route-policy frr
```

```
[SwitchD-rip-1] quit47
```

3. Verify the configuration:

Display route 4.4.4.4/32 on Switch S and you can view the backup next hop information.

```
[SwitchS] display ip routing-table 4.4.4.4 verbose
```

```
Routing Table : Public
```

```
Summary Count : 1
```

```
Destination: 4.4.4.4/32
  Protocol: RIP                Process ID: 1
  Preference: 100              Cost: 1
  IpPrecedence:                QoSLeId:
    NextHop: 13.13.13.2        Interface: vlan200
    BkNextHop: 12.12.12.2      BkInterface: vlan100
  RelyNextHop: 0.0.0.0         Neighbor : 0.0.0.0
  Tunnel ID: 0x0                Label: NULL
  BKTunnel ID: 0x0              BKLabel: NULL
  State: Active Adv            Age: 00h01m27s
  Tag: 0
```

Display route 1.1.1.1/32 on Switch D. You can find the backup next hop information.

```
[SwitchD] display ip routing-table 1.1.1.1 verbose
```

```
Routing Table : Public
```

```
Summary Count : 1
```

```
Destination: 1.1.1.1/32
  Protocol: RIP                Process ID: 1
  Preference: 100              Cost: 1
  IpPrecedence:                QoSLeId:
    NextHop: 13.13.13.1        Interface: vlan200
    BkNextHop: 24.24.24.2      BkInterface: vlan101
  RelyNextHop: 0.0.0.0         Neighbor : 0.0.0.0
  Tunnel ID: 0x0                Label: NULL
  BKTunnel ID: 0x0              BKLabel: NULL
  State: Active Adv            Age: 00h01m27s
  Tag: 0
```

Configuring BFD for RIP (single-hop detection in BFD echo packet mode)

Network requirements

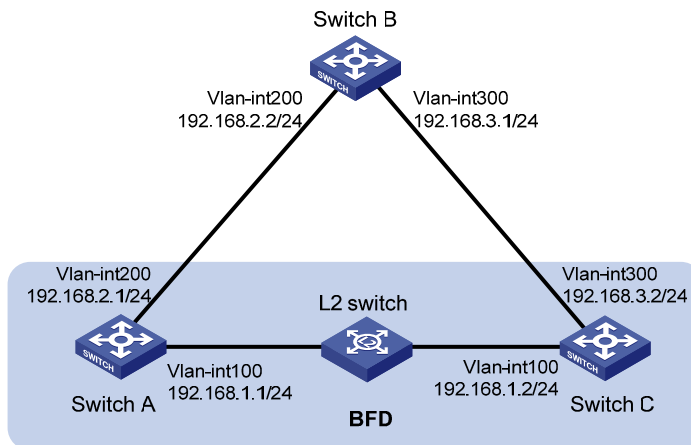
In the following figure, Switch A and Switch C are interconnected through a Layer 2 switch. VLAN-interface 100 of the two switches runs RIP process 1, BFD is enabled on VLAN-interface 100 of Switch A.

Switch A is connected to Switch C through Switch B. VLAN-interface 200 on Switch A runs RIP process 2; VLAN-interface 300 on Switch C, and VLAN-interface 200 and VLAN-interface 300 on Switch B run RIP process 1.

Configure a static route and enable static route redistribution into RIP on Switch C. Switch A learns the static route sent by Switch C, the outbound interface of the route is the interface connected to the Layer 2 switch.

Configure BFD so that when the link between Switch C and the Layer 2 switch fails, BFD can quickly detect the link failure and notify it to RIP, and the BFD session goes down. In response, RIP deletes the neighbor relationship with Switch C and the route information received from Switch C. Then, Switch A learns the static route sent by Switch C with the outbound interface being the interface connected to Switch B.

Figure 15 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure RIP basic functions:

Configure Switch A.

```
[SwitchA] rip 1
[SwitchA-rip-1] network 192.168.1.0
[SwitchA-rip-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] rip bfd enable
[SwitchA-Vlan-interface100] quit
[SwitchA] rip 2
[SwitchA-rip-2] network 192.168.2.0
```

Configure Switch B.

```
[SwitchB] rip 1
[SwitchB-rip-1] network 192.168.2.0
[SwitchB-rip-1] network 192.168.3.0
[SwitchB-rip-1] quit
```

Configure Switch C.

```
[SwitchC] rip 1
[SwitchC-rip-1] network 192.168.1.0
[SwitchC-rip-1] network 192.168.3.0
[SwitchC-rip-1] import-route static
[SwitchC-rip-1] quit
```

3. Configure BFD parameters:

Configure Switch A.

```
[SwitchA] bfd session init-mode active
[SwitchA] bfd echo-source-ip 11.11.11.11
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] bfd min-transmit-interval 500
[SwitchA-Vlan-interface100] bfd min-receive-interval 500
[SwitchA-Vlan-interface100] bfd detect-multiplier 7
[SwitchA-Vlan-interface100] quit
[SwitchA] quit
```

4. Configure a static route on Switch C:

```
[SwitchC] ip route-static 100.1.1.1 24 null 0
```

5. Verify the configuration:

Display the BFD session information of Switch A.

```
<SwitchA> display bfd session
Total Session Num: 1          Init Mode: Active
Session Working Under Echo Mode:
LD          SourceAddr      DestAddr      State   Holdtime   Interface
5           192.168.1.1      192.168.1.2  Up      2000ms     Vlan100
```

Display the RIP route 100.1.1.0/24 learned by Switch A.

```
<SwitchA> display ip routing-table 100.1.1.0 24 verbose
Routing Table : Public
Summary Count : 2
  Destination: 100.1.1.0/24
    Protocol: RIP          Process ID: 1
    Preference: 100        Cost: 1
    IpPrecedence:         QoSLeId:
      NextHop: 192.168.1.2  Interface: vlan-interface 100
      BkNextHop: 0.0.0.0    BkInterface:
      RelyNextHop: 0.0.0.0  Neighbor : 192.168.1.2
      Tunnel ID: 0x0        Label: NULL
      BKTunnel ID: 0x0      BKLabel: NULL
      State: Active Adv     Age: 00h00m47s
      Tag: 0
  Destination: 100.1.1.0/24
    Protocol: RIP          Process ID: 2
    Preference: 100        Cost: 2
    IpPrecedence:         QoSLeId:
      NextHop: 192.168.2.2  Interface: vlan-interface 200
      BkNextHop: 0.0.0.0    BkInterface:
      RelyNextHop: 0.0.0.0  Neighbor : 192.168.2.2
      Tunnel ID: 0x0        Label: NULL
      BKTunnel ID: 0x0      BKLabel: NULL
      State: Inactive Adv   Age: 00h12m50s
      Tag: 0
```

Enable RIP event debugging on Switch A.

```
<SwitchA> debugging rip 1 event
<SwitchA> terminal debugging
```

When the link between Switch C and the Layer 2 switch fails, Switch A can quickly detect the change.

```
%Jan 19 10:41:51:203 2008 SwitchA BFD/4/LOG:Sess[192.168.1.1/192.168.1.2,
Vlan-interface 100,Ctrl], Sta: UP->DOWN, Diag: 1
```

```
*Jan 19 10:33:12:813 2008 SwitchA RM/6/RMDEBUG: RIP-BFD: Message Type Disable, Connect
Type Direct-connect, Pkt Type Echo, Src IP Address 192.168.1.1, Src IFIndex4, Nbr IP
Address 192.168.1.2.
```

Display the BFD information of Switch A.

Switch A has deleted the neighbor relationship with Switch C and no output is displayed.

```
<SwitchA> display bfd session
```

Display the RIP routes of RIP process 1 on Switch A.

The RIP route learned from Switch C is no longer existent.

```
<SwitchA> display rip 1 route
```

```
Route Flags: R - RIP, T - TRIP
```

```
          P - Permanent, A - Aging, S - Suppressed, G - Garbage-collect
```

Display the RIP route 100.1.1.0/24 learned on Switch A.

```
<SwitchA> display ip routing-table 100.1.1.0 24 verbose
```

```
Routing Table : Public
```

```
Summary Count : 1
```

```
Destination: 100.1.1.0/24
```

```
  Protocol: RIP
```

```
  Process ID: 2
```

```
  Preference: 100
```

```
  Cost: 2
```

```
  IpPrecedence:
```

```
  QosLcId:
```

```
  NextHop: 192.168.2.2
```

```
  Interface: vlan-interface 200
```

```
  BkNextHop: 0.0.0.0
```

```
  BkInterface:
```

```
  RelyNextHop: 0.0.0.0
```

```
  Neighbor : 192.168.2.2
```

```
  Tunnel ID: 0x0
```

```
  Label: NULL
```

```
  BKTunnel ID: 0x0
```

```
  BKLabel: NULL
```

```
  State: Active Adv
```

```
  Age: 00h18m40s
```

```
  Tag: 0
```

Configuring BFD for RIP (bidirectional detection in BFD control packet mode)

Network requirements

In the following figure, Switch A is connected to Switch C through Switch B. VLAN-interface 100 on Switch A, VLAN-interface 200 on Switch C, and VLAN-interface 200 and VLAN-interface 100 on Switch B run RIP process 1.

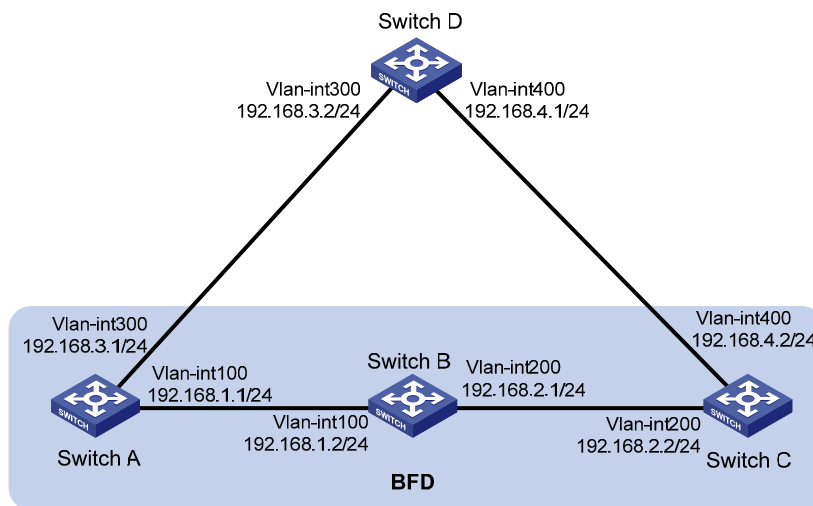
Configure a static route to Switch C on Switch A, and configure a static route to Switch A on Switch C. Enable BFD on VLAN-interface 100 of Switch A and VLAN-interface 200 of Switch C.

Switch A is connected to Switch C through Switch D. VLAN-interface 300 on Switch A runs RIP process 2; VLAN-interface 400 on Switch C, and VLAN-interface 300 and VLAN-interface 400 on Switch D run RIP process 1.

Enable static route redistribution into RIP on Switch A and Switch C so that Switch A and Switch C have routes to send to each other. Switch A learns the static route sent by Switch C, the outbound interface is the interface connected to Switch B.

Configure BFD so that when the link between Switch B and Switch C fails, BFD can quickly detect the link failure and notify it to RIP, and the BFD session goes down. In response, RIP deletes the neighbor relationship with Switch C and the route information received from Switch C. Then, Switch A learns the static route sent by Switch C, the outbound interface of the route is the interface connected to Switch D.

Figure 16 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure RIP basic functions and enable static route redistribution so that Switch A and Switch C have routes to send to each other:

Configure Switch A.

```
[SwitchA] rip 1
[SwitchA-rip-1] network 192.168.1.0
[SwitchA-rip-1] peer 192.168.2.2
[SwitchA-rip-1] undo validate-source-address
[SwitchA-rip-1] import-route static
[SwitchA-rip-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] rip bfd enable
[SwitchA-Vlan-interface100] quit
[SwitchA] rip 2
[SwitchA-rip-2] network 192.168.3.0
[SwitchA-rip-2] quit
```

Configure Switch C.

```
[SwitchC] rip 1
[SwitchC-rip-1] network 192.168.2.0
[SwitchC-rip-1] network 192.168.4.0
[SwitchC-rip-1] peer 192.168.1.1
[SwitchC-rip-1] undo validate-source-address
[SwitchC-rip-1] import-route static
```

```
[SwitchC-rip-1] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] rip bfd enable
[SwitchC-Vlan-interface200] quit
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] rip 1
[SwitchD-rip-1] network 192.168.3.0
[SwitchD-rip-1] network 192.168.4.0
```

3. Configure BFD parameters:

Configure Switch A.

```
[SwitchA] bfd session init-mode active
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] bfd min-transmit-interval 500
[SwitchA-Vlan-interface100] bfd min-receive-interval 500
[SwitchA-Vlan-interface100] bfd detect-multiplier 7
[SwitchA-Vlan-interface100] quit
```

Configure Switch C.

```
[SwitchC] bfd session init-mode active
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] bfd min-transmit-interval 500
[SwitchC-Vlan-interface200] bfd min-receive-interval 500
[SwitchC-Vlan-interface200] bfd detect-multiplier 7
[SwitchC-Vlan-interface200] quit
```

4. Configure static routes.

Configure a static route to Switch C on Switch A.

```
[SwitchA] ip route-static 192.168.2.0 24 vlan-interface 100 192.168.1.2
[SwitchA] ip route-static 101.1.1.0 24 null 0
[SwitchA] quit
```

Configure a static route to Switch A on Switch C.

```
[SwitchC] ip route-static 192.168.1.0 24 vlan-interface 200 192.168.2.1
[SwitchC] ip route-static 100.1.1.0 24 null 0
```

! IMPORTANT:

If you specify null 0 interface as the output interface for a static route, do not specify the IP address of a directly connected network as the destination IP address.

5. Verify the configuration:

Display the BFD session information of Switch A.

```
<SwitchA> display bfd session
Total Session Num: 1          Init Mode: Active
Session Working Under Ctrl Mode:
LD/RD      SourceAddr      DestAddr      State Holdtime Interface
6/3        192.168.1.1     192.168.2.2   Up    1700ms   vlan100
```

Display the RIP route 100.1.1.0/24 learned on Switch A.

```
<SwitchA> display ip routing-table 100.1.1.0 24 verbose
Routing Table : Public
```

```

Summary Count : 2
  Destination: 100.1.1.0/24
    Protocol: RIP          Process ID: 1
    Preference: 100        Cost: 1
  IpPrecedence:          QoSLeId:
    NextHop: 192.168.1.2   Interface: vlan-interface 100
    BkNextHop: 0.0.0.0     BkInterface:
  RelyNextHop: 0.0.0.0    Neighbor : 192.168.1.2
    Tunnel ID: 0x0         Label: NULL
  BKTunnel ID: 0x0       BKLabel: NULL
    State: Active Adv      Age: 00h00m47s
    Tag: 0
  Destination: 100.1.1.0/24
    Protocol: RIP          Process ID: 2
    Preference: 100        Cost: 2
  IpPrecedence:          QoSLeId:
    NextHop: 192.168.3.2   Interface: vlan-interface 300
    BkNextHop: 0.0.0.0     BkInterface:
  RelyNextHop: 0.0.0.0    Neighbor : 192.168.3.2
    Tunnel ID: 0x0         Label: NULL
  BKTunnel ID: 0x0       BKLabel: NULL
    State: Inactive Adv    Age: 00h12m50s
    Tag: 0

```

Enable RIP event debugging on Switch A.

```

<SwitchA> debugging rip 1 event
<SwitchA> terminal debugging

```

When the link between Switch B and Switch C fails, Switch A quickly detects the link state change.

```

%Jan 19 10:41:51:203 2008 SwitchA BFD/4/LOG:Sess[192.168.1.1/192.168.2.2,
Vlan-interface 100, Ctrl], Sta: UP->DOWN, Diag: 1
*Jan 19 10:41:51:203 2008 SwitchA RM/6/RMDEBUG: RIP-BFD: Message Type Disable, Connect
Type Indirect-connect, Pkt Type Control, Src IP Address 192.168.1.1, Src IFIndex 4,
Nbr IP Address 192.168.2.2.

```

Display the BFD information of Switch A.

Switch A has deleted the neighbor relationship with Switch C and no output is displayed.

```

<SwitchA> display bfd session

```

Display the RIP routes of RIP process 1 on Switch A.

The RIP route learned from Switch C is no longer existent.

```

<SwitchA> display rip 1 route
Route Flags: R - RIP, T - TRIP
              P - Permanent, A - Aging, S - Suppressed, G - Garbage-collect
-----

```

Display the RIP route 100.1.1.0/24 learned on Switch A.

```

<SwitchA> display ip routing-table 100.1.1.0 24 verbose
Routing Table : Public
Summary Count : 1

```

```

Destination: 100.1.1.0/24
  Protocol: RIP                Process ID: 2
  Preference: 100              Cost: 2
IpPrecedence:                  QoSLeId:
  NextHop: 192.168.3.2        Interface: vlan-interface 300
  BkNextHop: 0.0.0.0          BkInterface:
  RelyNextHop: 0.0.0.0        Neighbor : 192.168.3.2
  Tunnel ID: 0x0              Label: NULL
  BKTunnel ID: 0x0            BKLabel: NULL
  State: Active Adv           Age: 00h18m40s
  Tag: 0

```

Troubleshooting RIP

No RIP updates received

Symptom

No RIP updates are received when the links function.

Analysis

After enabling RIP, you must use the **network** command to enable corresponding interfaces. Ensure no interfaces are disabled from handling RIP messages.

If the peer is configured to send multicast messages, the same must be configured on the local end.

Solution

1. Use the **display current-configuration** command to check RIP configuration.
2. Use the **display rip** command to check whether an interface is disabled.

Route oscillation occurred

Symptom

When all links function, route oscillation occurs on the RIP network. After displaying the routing table, you may find some routes intermittently appear and disappear in the routing table.

Analysis

In the RIP network, make sure that all the same timers within the entire network are identical and have logical relationships between them. For example, the timeout timer value must be greater than the update timer value.

Solution

1. Use the **display rip** command to check the configuration of RIP timers.
2. Use the **timers** command to adjust timers properly.

Configuring OSPF

Hardware compatibility

The HP 5500 SI Switch Series does not support OSPF.

Introduction to OSPF

Open Shortest Path First (OSPF) is a link state interior gateway protocol developed by the OSPF working group of the IETF. Now, OSPF version 2 (RFC 2328) is used. Unless otherwise noted, OSPF refers to OSPFv2 throughout this chapter. The term "router" in this chapter refers to both routers and Layer 3 switches.

OSPF has the following features:

- **Wide scope**—Supports various network sizes and up to several hundred routers in an OSPF routing domain.
- **Fast convergence**—Transmits routing updates instantly upon network topology changes.
- **Loop-free**—Computes routes with the shortest path first (SPF) algorithm to avoid routing loops.
- **Area-based network partition**—Splits an AS into different areas to facilitate management. In addition, routing information transmitted between areas is summarized to reduce traffic and routing table sizes.
- **Equal-cost multi-path (ECMP) routing**—Supports multiple equal-cost routes to a destination.
- **Routing hierarchy**—Supports a four-level routing hierarchy that prioritizes routes into intra-area, inter-area, external Type-1, and external Type-2 routes.
- **Authentication**—Supports interface-based packet authentication to ensure the security of packet exchange.
- **Support for multicast**—Multicasts protocol packets on some types of links.

Basic concepts

Autonomous System

An Autonomous System (AS) comprises a group of routers that run the same routing protocol.

OSPF route computation

OSPF computes routes in an area as follows:

- Based on the network topology around itself, each router generates Link State Advertisements (LSAs) and sends them to other routers in update packets.
- Each OSPF router collects LSAs from other routers to compose a link state database (LSDB). An LSA describes the network topology around a router, and the LSDB describes the entire network topology of the AS.
- Each router transforms the LSDB in the area to a weighted directed graph, which is the topology of the entire network. All the routers of the area have the same graph.

- Each router uses the SPF algorithm to compute a shortest path tree showing the routes to the nodes in the AS. The router itself is the root of the tree.

Router ID

An OSPF process running on a router must have its own router ID. This ID is a 32-bit unsigned integer that uniquely identifies the router in the AS.

OSPF packets

OSPF uses the following types of packets:

- **Hello**—Periodically sent to find and maintain neighbors, containing the values of some timers, information about the DR, BDR, and known neighbors.
- **Database description (DD)**—Describes the digest of each LSA in the LSDB, exchanged between two routers for data synchronization.
- **Link state request (LSR)**—Requests needed LSAs from the neighbor. After exchanging the DD packets, the two routers know which LSAs of the neighbor are missing from their LSDBs. They then send an LSR packet to each other, requesting the missing LSAs. The LSA packet contains the digest of the missing LSAs.
- **Link state update (LSU)**—Transmits the requested LSAs to the neighbor.
- **Link state acknowledgment (LSAck)**—Acknowledges received LSU packets. It contains the headers of received LSAs (an LSAck packet can acknowledge multiple LSAs).

LSA types

OSPF sends routing information in LSAs, which—as defined in RFC 2328—have the following types:

- **Router LSA**—Type-1 LSA, originated by all routers, flooded throughout a single area only. This LSA describes the collected states of the router's interfaces to an area.
- **Network LSA**—Type-2 LSA, originated for broadcast and NBMA networks by the designated router, flooded throughout a single area only. This LSA contains the list of routers connected to the network.
- **Network Summary LSA**—Type-3 LSA, originated by ABRs (Area Border Routers), and flooded throughout the LSA's associated area. Each summary-LSA describes a route to a destination outside the area, yet still inside the AS (an inter-area route).
- **ASBR Summary LSA**—Type-4 LSA, originated by ABRs and flooded throughout the LSA's associated area. Type 4 summary-LSAs describe routes to ASBR (Autonomous System Boundary Router).
- **AS External LSA**—Type-5 LSA, originated by ASBRs, and flooded throughout the AS (except stub and NSSA areas). Each AS-external-LSA describes a route to another AS.
- **NSSA LSA**—Type-7 LSA, as defined in RFC 1587, originated by ASBRs in NSSAs (Not-So-Stubby Areas) and flooded throughout a single NSSA. NSSA LSAs describe routes to other ASs.
- **Opaque LSA**—A proposed type of LSA, the format consisting of a standard LSA header and application specific information. Opaque LSAs are used by the OSPF protocol or by some application to distribute information into the OSPF routing domain. The opaque LSA includes Type 9, Type 10, and Type 11. The Type 9 opaque LSA is flooded into the local subnet, the Type 10 is flooded into the local area, and the Type 11 is flooded throughout the AS.

Neighbor and Adjacency

In OSPF, "neighbor" and "adjacency" are different concepts.

- **Neighbor**—After startup, OSPF sends a hello packet on each OSPF interface. A router that receives the hello packet checks parameters in the packet. If the parameters match its own, the router considers the sending router an OSPF neighbor.
- **Adjacency**—Two OSPF neighbors establish an adjacency relationship to synchronize their LSDBs. Therefore, any two neighbors without exchanging route information do not establish an adjacency.

Area based OSPF network partition

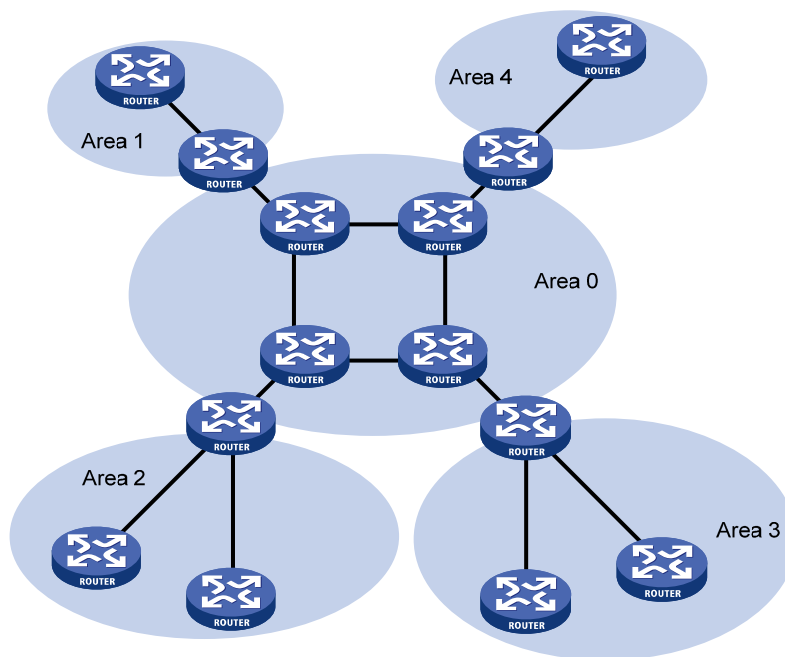
Network partition

In a large OSPF routing domain, the LSDB becomes very huge and SPF computation consumes many storage and CPU resources.

In addition, because topology changes can easily occur, OSPF packets generated for route information synchronization are enormous, occupying excessive bandwidth.

To solve these problems, OSPF splits an AS into multiple areas, each of which is identified by an area ID. The boundaries between areas are routers rather than links. A network segment (or a link) can only reside in one area. An OSPF interface must be specified to belong to its attached area, as shown in Figure 17.

Figure 17 Area based OSPF network partition



After network partition, ABRs perform route summarization to reduce the number of LSAs advertised to other areas and minimize the effect of topology changes.

Backbone area and virtual links

Each AS has a backbone area that distributes routing information between non-backbone areas. Routing information between non-backbone areas must be forwarded by the backbone area. OSPF requires the following:

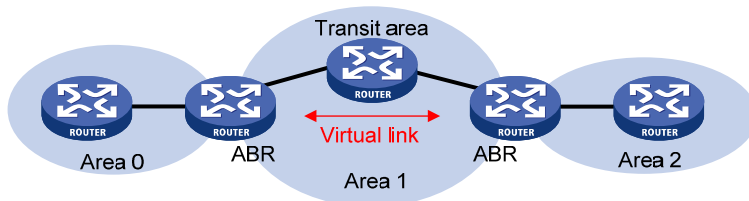
- All non-backbone areas must maintain connectivity to the backbone area.
- The backbone area itself must maintain connectivity.

In practice, the requirements may not be satisfied due to lack of physical links. OSPF virtual links can solve this problem.

A virtual link is established between two ABRs through a non-backbone area and is configured on both ABRs to take effect. The non-backbone area is called a transit area.

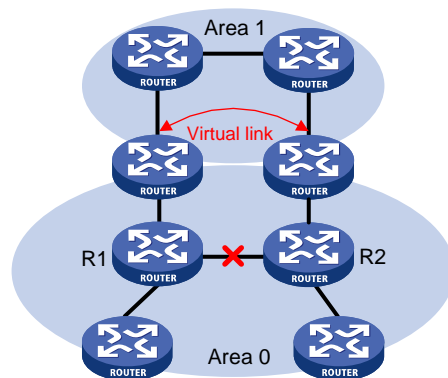
In the following figure, Area 2 has no direct physical link to the backbone area 0. You can configure a virtual link between the two ABRs to connect Area 2 to the backbone area.

Figure 18 Virtual link application 1



Virtual links can also be used to provide redundant links. If the backbone area cannot maintain internal connectivity due to the failure of a physical link, you can configure a virtual link to replace the failed physical link, as shown in Figure 19.

Figure 19 Virtual link application 2



The virtual link between the two ABRs acts as a point-to-point connection. You can configure interface parameters such as hello interval on the virtual link as they are configured on a physical interface.

The two ABRs on the virtual link unicast OSPF packets to each other, and the OSPF routers in between convey these OSPF packets as normal IP packets.

Stub area

A stub area does not distribute Type-5 LSAs, so the routing table size and amount of routing information in this area are reduced significantly. The ABR generates a default route into the area.

You can configure the stub area as a totally stub area, where the ABR advertises neither inter-area routes nor external routes.

Stub area configuration is optional, and not every area is eligible to be a stub area. In general, a stub area resides on the border of the AS.

When you configure a totally stub area, follow these guidelines:

- The backbone area cannot be a totally stub area.
- To configure an area as a stub area, the **stub** command must be configured on routers in the area.

- To configure an area as a totally stub area, the **stub** command must be configured on routers in the area, and the ABR of the area must be configured with the **stub [no-summary]** command.
- A totally stub area cannot have an ASBR because AS external routes cannot be distributed into the stub area.
- Virtual links cannot transit totally stub areas.

NSSA area

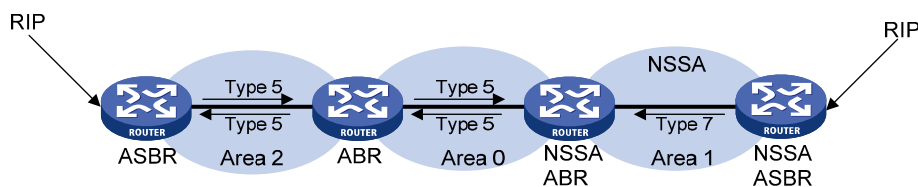
Similar to a stub area, an NSSA area does not import AS external LSAs (Type-5 LSAs), but can import Type-7 LSAs generated by the NSSA ASBR. The NSSA ABR translates Type-7 LSAs into Type-5 LSAs and advertises the Type-5 LSAs to other areas.

In the following figure, the OSPF AS contains Area 1, Area 2, and Area 0. The other two ASs run RIP. Area 1 is an NSSA area. The ASBR redistributes RIP routes in Type-7 LSAs into Area 1. Upon receiving these Type-7 LSAs, the NSSA ABR translates them to Type-5 LSAs, and then advertises the Type-5 LSAs to Area 0.

The ASBR of Area 2 redistributes RIP routes in Type-5 LSAs into the OSPF routing domain. However, Area 1 does not receive these Type-5 LSAs because it is an NSSA area.

Virtual links cannot transit NSSA areas.

Figure 20 NSSA area



Comparison between the areas

Figure 21 Comparison between the areas

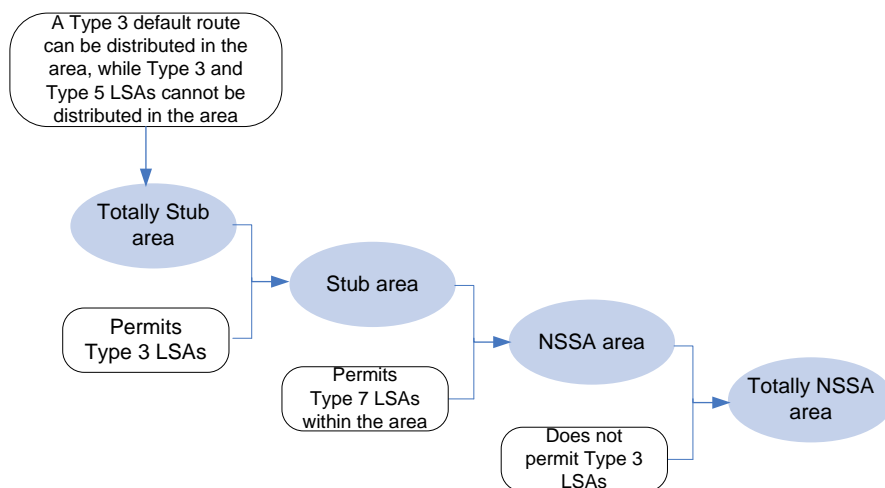


Figure 21 shows the comparison of the areas:

- In a totally stub area, the ABR distributes a Type 3 default route, rather than external routes and inter-area routes.
- A stub area can import inter-area routes, but a stub area cannot.

- An NSSA area can import external routes in Type 7 LSAs through the ASBR, but a stub area cannot.
- A totally NSSA area cannot import inter-area routes but an NSSA area can.

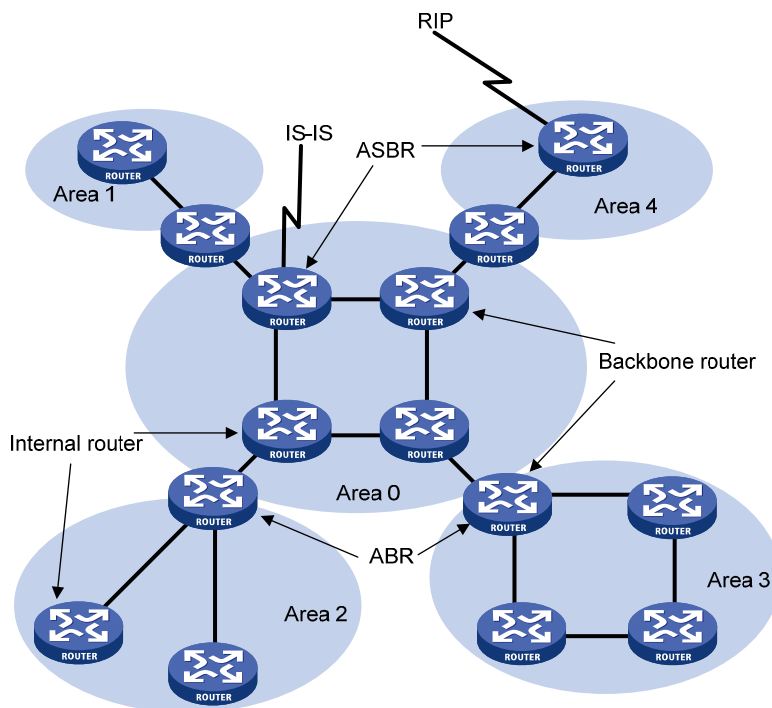
Router types

Router classification

The following are OSPF router types and their positions in the AS:

- **Internal router**—All interfaces on an internal router belong to one OSPF area.
- **Area Border Router (ABR)**—An ABR belongs to more than two areas, one of which must be the backbone area. It connects the backbone area to a non-backbone area. The connection between an ABR and the backbone area can be physical or logical.
- **Backbone router**—At least one interface of a backbone router must reside in the backbone area. All ABRs and internal routers in area 0 are backbone routers.
- **Autonomous System Boundary Router (ASBR)**—A router exchanging routing information with another AS is an ASBR, which may not reside on the border of the AS. It can be an internal router or an ABR.

Figure 22 OSPF router types



Route types

OSPF prioritize routes into the following levels:

- Intra-area route
- Inter-area route
- Type-1 external route
- Type-2 external route

The intra-area and inter-area routes describe the network topology of the AS. The external routes describe routes to external ASs.

OSPF classifies external routes as Type-1 or Type-2. A Type-1 external route has high credibility. The cost from a router to the destination of the Type-1 external route = the cost from the router to the corresponding ASBR + the cost from the ASBR to the destination of the external route.

A Type-2 external route has low credibility, so OSPF considers the cost from the ASBR to the destination of the Type-2 external route is much greater than the cost from the ASBR to an OSPF internal router. The cost from the internal router to the destination of the Type-2 external route = the cost from the ASBR to the destination of the Type-2 external route. If two routes to the same destination have the same cost, OSPF takes the cost from the router to the ASBR into consideration to determine the best route.

OSPF network classification

OSPF network types

OSPF classifies networks into the following types depending on different link layer protocols:

- **Broadcast**—When the link layer protocol is Ethernet or FDDI, OSPF considers the network type as broadcast by default. On a broadcast network, hello, LSU, and LSAck packets are multicast to 224.0.0.5 that identifies all OSPF routers or 224.0.0.6 that identifies the DR, and DD packets and LSR packets are unicast.
- **NBMA (Non-Broadcast Multi-Access)**—When the link layer protocol is Frame Relay, ATM, or X.25, OSPF considers the network type as NBMA by default. OSPF packets are unicast on a NBMA network.
- **P2MP (point-to-multipoint)**—By default, OSPF considers no link layer protocol as P2MP, which is a conversion from other network types such as NBMA. On a P2MP network, OSPF packets are multicast to 224.0.0.5.
- **P2P (point-to-point)**—When the link layer protocol is PPP or HDLC, OSPF considers the network type as P2P. On a P2P network, OSPF packets are multicast to 224.0.0.5.

NBMA network configuration guidelines

Typical NBMA networks include ATM and Frame Relay networks.

Because NBMA interfaces cannot broadcast hello packets, you must specify neighbors manually and configure router priorities for the neighbors.

An NBMA network is fully meshed, which means any two routers in the NBMA network have a direct virtual circuit for communication. If direct connections are not available between some routers, the network type of associated interfaces must be configured as P2MP. If such an interface has only one neighbor, configure its network type as P2P.

The differences between NBMA and P2MP networks are as follows:

- NBMA networks are fully meshed, non-broadcast, and multi access. P2MP networks are not required to be fully meshed.
On an NBMA network, you must elect the DR and BDR, while on a P2MP network, DR and BDR are not available.
- NBMA is the default network type, but P2MP is a conversion from another network type, such as NBMA.
- On a NBMA network, OSPF packets are unicast, and neighbors are manually configured on routers. On a P2MP network, OSPF packets are multicast.

DR and BDR

Introduction

On a broadcast or NBMA network, any two routers need to establish an adjacency to exchange routing information with each other. If n routers are present on the network, $n(n-1)/2$ adjacencies are required. In addition, any topology change on the network results in traffic for route synchronization, which consumes many system and bandwidth resources. The Designated Router (DR) was introduced to solve this problem. On a network, a DR is elected to advertise routing information among other routers.

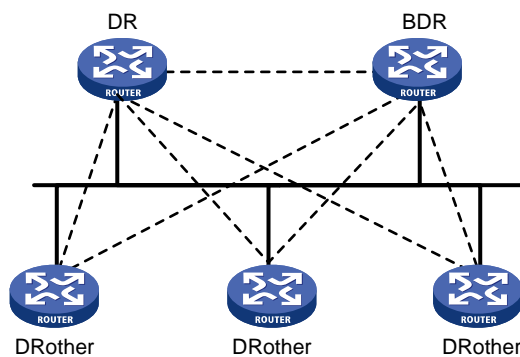
If the DR fails, routers on the network have to elect another DR and synchronize information with the new DR. It is time-consuming and prone to routing calculation errors. The Backup Designated Router (BDR) can solve this problem.

The BDR is elected along with the DR and establishes adjacencies with all other routers. When the DR fails, the BDR becomes the new DR in a very short time. Meanwhile, other routers elect a new BDR.

Routers other than the DR and BDR are called "DRothers". They do not establish adjacencies with one another. Thus the number of adjacencies is reduced.

In [Figure 23](#), solid lines are Ethernet physical links, and dashed lines represent OSPF adjacencies. In the network with the DR and BDR, only seven adjacencies are needed.

Figure 23 DR and BDR in a network



DR and BDR election

Routers in a network elect the DR and BDR according to their router priorities and router IDs. Routers with a router priority value higher than 0 are candidates for DR/BDR election.

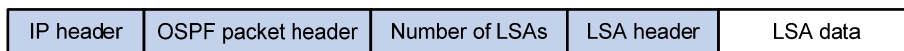
The election votes are hello packets. Each router sends the DR elected by itself in a hello packet to all the other routers. If two routers on the network declare themselves as the DR, the router with the higher router priority wins. If router priorities are the same, the router with the higher router ID wins. In addition, a router with router priority 0 cannot become the DR or BDR.

- DR election is available on broadcast and NBMA interfaces rather than P2P and P2MP interfaces.
- A DR is an interface of a router and belongs to a single network segment. Another interface of the router may be a BDR or DROther.
- If a router with the highest router priority is added after DR/BDR election, the router cannot become the DR immediately.
- The DR may not be the router with the highest priority in a network, and the BDR may not be the router with the second highest priority.

OSPF packet formats

OSPF packets are directly encapsulated into IP packets. OSPF uses the IP protocol number 89. The format of an OSPF LSU packet is shown in [Figure 24](#).

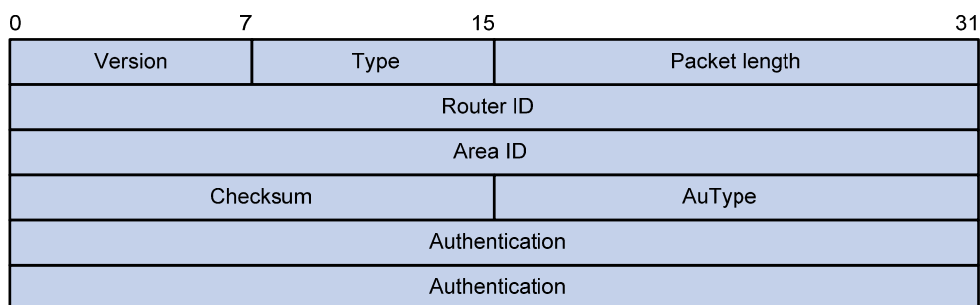
Figure 24 OSPF packet format



OSPF packet header

OSPF packets are classified into five types that have the same packet header.

Figure 25 OSPF packet header



Major fields of the OSPF packet header are as follows:

- **Version**—OSPF version number, which is 2 for OSPFv2.
- **Type**—OSPF packet type from 1 to 5, corresponding to hello, DD, LSR, LSU, and LSAck, respectively.
- **Packet length**—Total length of the OSPF packet in bytes, including the header.
- **Router ID**—ID of the advertising router.
- **Area ID**—ID of the area where the advertising router resides.
- **Checksum**—Checksum of the message.
- **AuType**—Authentication type, ranging from 0 to 2, corresponding to non-authentication, simple (plaintext) authentication, and MD5 authentication, respectively.
- **Authentication**—Information determined by authentication type. It is not defined for authentication type 0. It is defined as password information for authentication type 1, and defined as Key ID, MD5 authentication data length, and sequence number for authentication type 2.

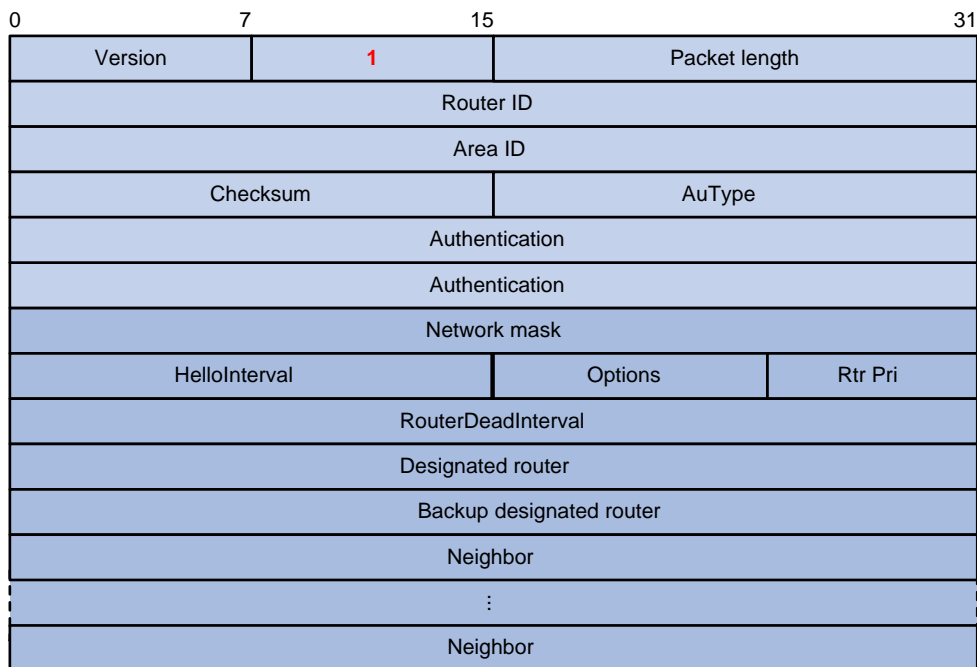
NOTE:

MD5 authentication data is added following an OSPF packet rather than contained in the Authentication field.

Hello packet

A router sends hello packets periodically to find and maintain neighbor relationships, and to elect the DR or BDR, including information about values of timers, DR, BDR, and neighbors that are already known.

Figure 26 Hello packet format



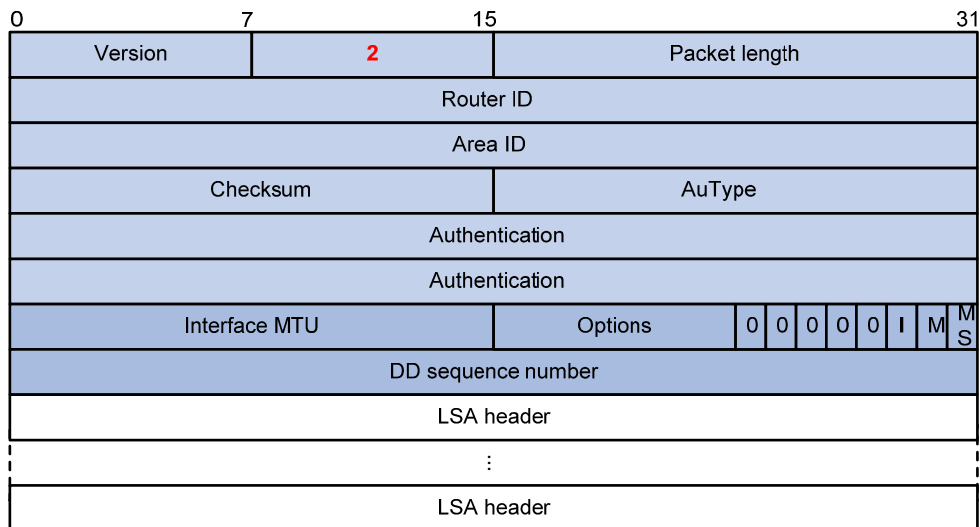
Major fields of the hello packet are as follows:

- **Network mask**—Network mask associated with the router’s sending interface. If two routers have different network masks, they cannot become neighbors.
- **HelloInterval**—Interval for sending hello packets. If two routers have different intervals, they cannot become neighbors.
- **Rtr Pri**—Router priority. A value of 0 means the router cannot become the DR or BDR.
- **RouterDeadInterval**—Time before declaring a silent router down. If two routers have different dead intervals, they cannot become neighbors.
- **Designated router**—IP address of the DR.
- **Backup designated router**—IP address of the BDR.
- **Neighbor**—Router ID of the neighbor router.

DD packet

Two routers exchange database description (DD) packets, describing their LSDBs for database synchronization. A DD packet contains only the headers of LSAs to reduce traffic.

Figure 27 DD packet format



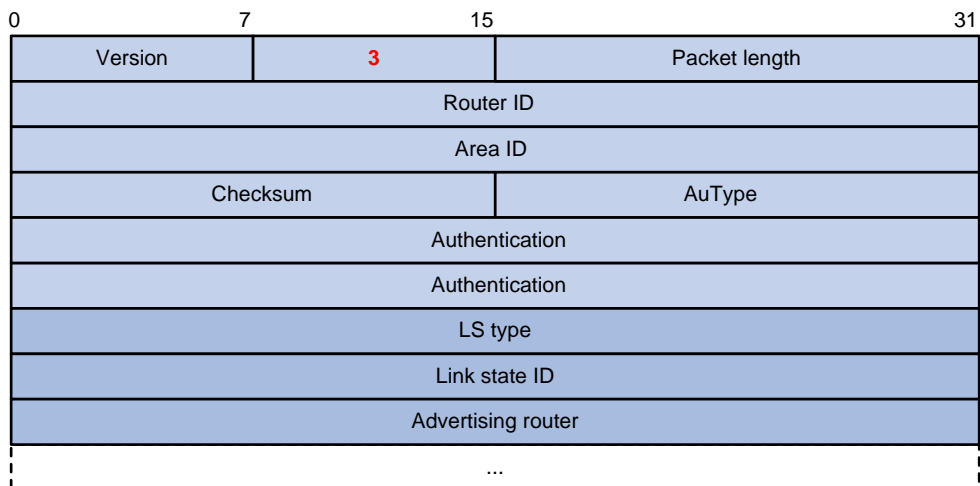
Major fields of the DD packets are as follows:

- **Interface MTU**—Specifies the largest IP datagram in bytes that the interface can send without fragmentation.
- **I (Initial)**—The Init bit, which is set to 1 if the packet is the first DD packet. It is set to 0 if not.
- **M (More)**—The More bit, which is set to 0 if the packet is the last DD packet. It is set to 1 if more DD packets are to follow.
- **MS (Master/Slave)**—The Master/Slave bit. When set to 1, it indicates that the router is the master during the database exchange process; otherwise, the router is the slave router.
- **DD sequence number**—Used to sequence the collection of DD packets. The initial value is set by the master. The DD sequence number then increments until the complete database description has been sent.

LSR packet

After exchanging DD packets, two routers know which LSAs of the peer are missing from the local LSDB. Then, they send (link state request) LSR packets to request the missing LSAs. An LSR packet contains the brief of the missing LSAs.

Figure 28 LSR packet format



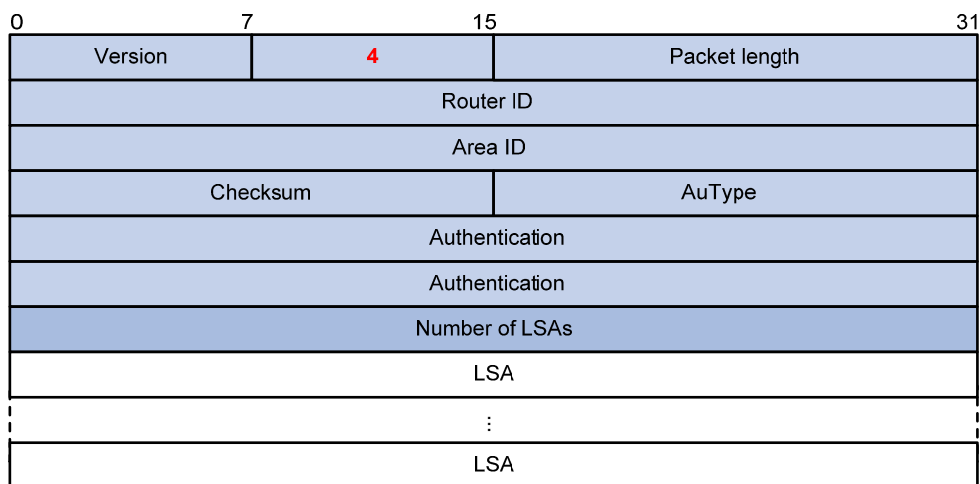
Major fields of the LSR packets are as follows:

- **LS type**—Type of the LSA to be requested. Type 1 for example indicates the Router LSA.
- **Link state ID**—Determined by LSA type.
- **Advertising router**—ID of the router that sent the LSA.

LSU packet

LSU (Link State Update) packets are used to send the requested LSAs to the peer. Each packet carries a collection of LSAs.

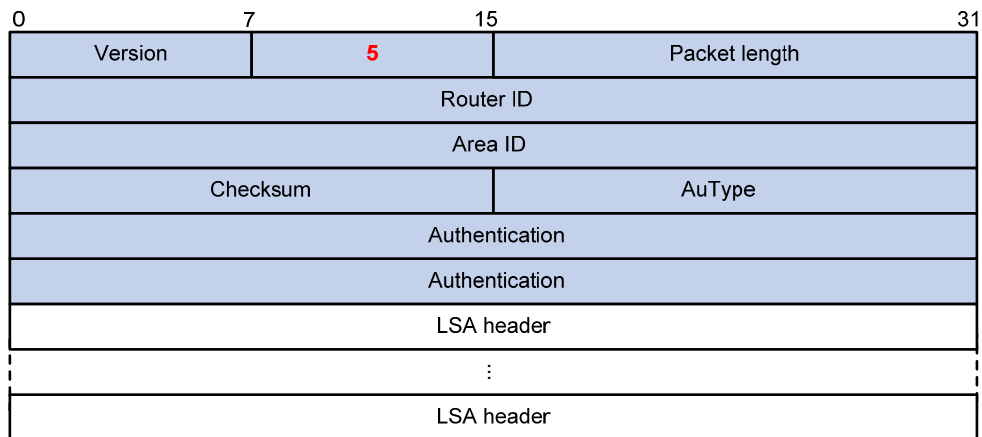
Figure 29 LSU packet format



LSAck packet

Link State Acknowledgment (LSAck) packets are used to acknowledge received LSU packets. An LSAck packet carries the headers of LSAs to be acknowledged.

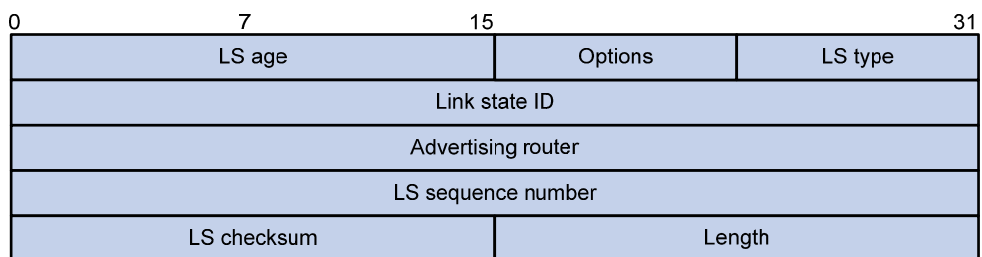
Figure 30 LSAck packet format



LSA header format

All LSAs have the same header.

Figure 31 LSA header format



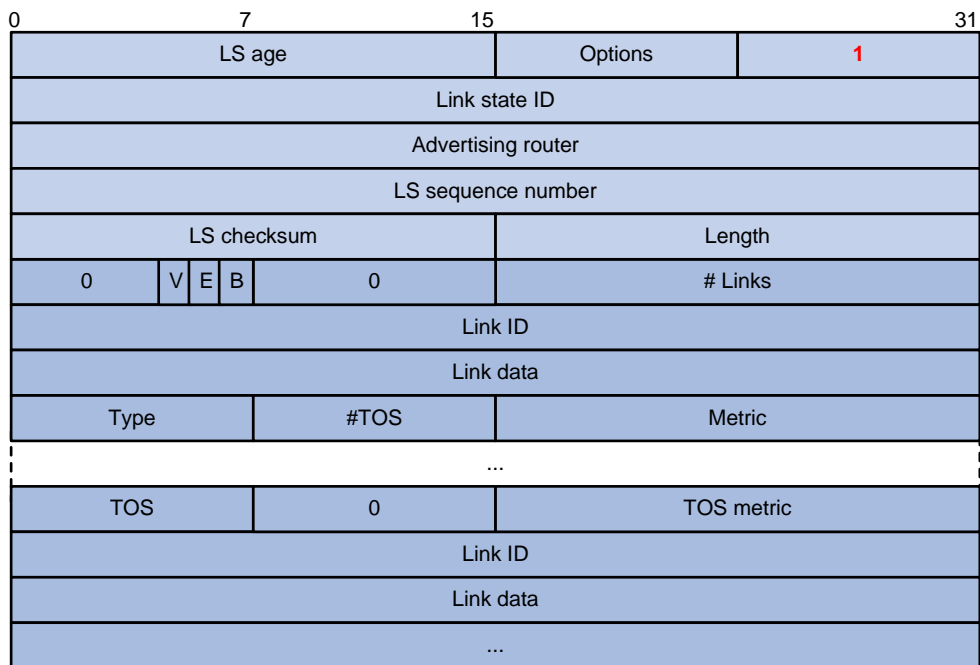
Major fields of the LSA header are as follows:

- **LS age**—Time, in seconds, elapsed since the LSA was originated. An LSA ages in the LSDB (added by 1 per second), but does not age during transmission.
- **LS type**—Type of the LSA.
- **Link state ID**—The contents of this field depend on the LSA's type.
- **LS sequence number**—Used by other routers to judge new and old LSAs.
- **LS checksum**—Checksum of the LSA except the LS age field.
- **Length**—Length in bytes of the LSA, including the LSA header.

LSAs formats

- Router LSA

Figure 32 Router LSA format



Major fields of the Router LSA are as follows:

- **Link state ID**—ID of the router that originated the LSA.
- **V (Virtual Link)**—Set to 1 if the router that originated the LSA is a virtual link endpoint.
- **E (External)**—Set to 1 if the router that originated the LSA is an ASBR.
- **B (Border)**—Set to 1 if the router that originated the LSA is an ABR.
- **# Links**—Number of router links (interfaces) to the area, as described in the LSA.
- **Link ID**—Determined by link type.
- **Link data**—Determined by link type.
- **Type**—Link type. A value of 1 indicates a point-to-point link to a remote router; a value of 2 indicates a link to a transit network; a value of 3 indicates a link to a stub network; and a value of 4 indicates a virtual link.
- **#TOS**—Number of different TOS metrics given for this link. If no TOS metric is given for the link, this field is set to 0. TOS is not supported in RFC 2328. The #TOS field is reserved for early versions of OSPF.
- **Metric**—Cost of using this router link.
- **TOS**—IP Type of Service that this metric refers to.
- **TOS metric**—TOS-specific metric information.
- Network LSA

A Network LSA is originated by the DR on a broadcast or NBMA network. The LSA describes all routers attached to the network.

Figure 33 Network LSA format

0	7	15	31
LS age		Options	2
Link state ID			
Advertising router			
LS sequence number			
LS checksum		Length	
Network mask			
Attached router			
...			

Major fields of the Network LSA are as follows:

- **Link state ID**—The interface address of the DR.
- **Network mask**—The mask of the network (a broadcast or NBMA network).
- **Attached router**—The IDs of the routers, which are adjacent to the DR, including the DR itself.
- Summary LSA

Network summary LSAs (Type-3 LSAs) and ASBR summary LSAs (Type-4 LSAs) are originated by ABRs. Except for the Link state ID field, the formats of Type 3 and 4 summary-LSAs are identical.

Figure 34 Summary LSA format

0	7	15	31
LS age		Options	3or4
Link state ID			
Advertising router			
LS sequence number			
LS checksum		Length	
Network mask			
0	Metric		
TOS	TOS metric		
...			

Major fields of the Summary LSA are as follows:

- **Link state ID**—For a Type-3 LSA, it is an IP address outside the area. For a type 4 LSA, it is the router ID of an ASBR outside the area.
- **Network mask**—The network mask for the type 3 LSA. It is set to 0.0.0.0 for the Type-4 LSA.
- **Metric**—The metric to the destination.

NOTE:

A Type-3 LSA can be used to advertise a default route if the link state ID and network mask are set to 0.0.0.0.

- AS external LSA

An AS external LSA is originated by an ASBR, and describes routing information to a destination outside the AS.

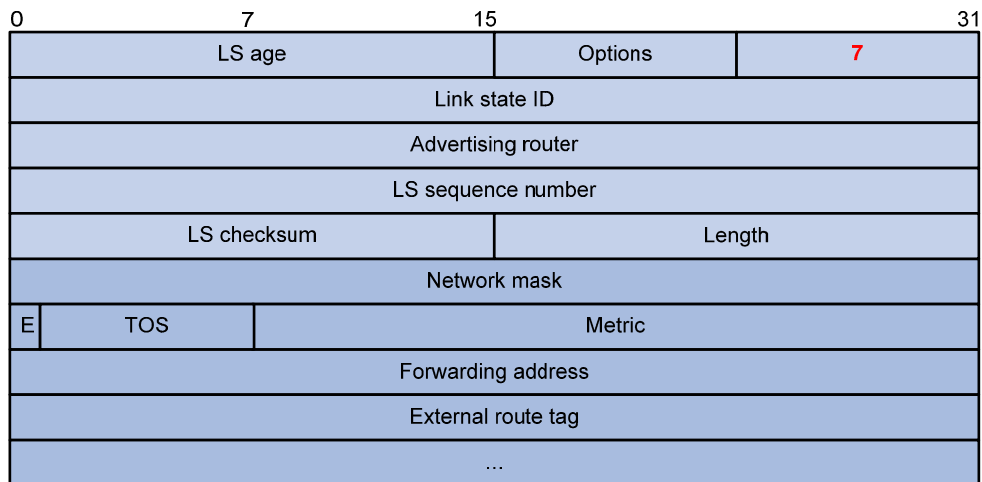
Figure 35 AS external LSA format

0	7	15	31
LS age		Options	5
Link state ID			
Advertising router			
LS sequence number			
LS checksum		Length	
Network mask			
E	0	Metric	
Forwarding address			
External route tag			
E	TOS	TOS metric	
Forwarding address			
External route tag			
...			

Major fields of the AS external LSA are as follows:

- **Link state ID**—The IP address of another AS to be advertised. When describing a default route, the Link state ID is always set to default destination (0.0.0.0) and the network mask is set to 0.0.0.0
- **Network mask**—The IP address mask for the advertised destination
- **E (External Metric)**—The type of the external metric value, which is set to 1 for type 2 external routes, and set to 0 for type 1 external routes. See "[Route types](#)" for a description of external route types.
- **Metric**—The metric to the destination.
- **Forwarding address**—Data traffic for the advertised destination is forwarded to this address.
- **External route tag**—A tag attached to each external route. This is not used by the OSPF protocol. It may be used to manage external routes.
- **NSSA external LSA**
An NSSA external LSA originates from the ASBR in an NSSA, and is flooded in the NSSA area only. It has the same format as the AS external LSA.

Figure 36 NSSA external LSA format



Supported features

Multi-process

This feature allows multiple OSPF processes to run on a router both simultaneously and independently. Routing information interactions between different processes simulate interactions between different routing protocols. Multiple OSPF processes can use the same RID.

An interface of a router can only belong to a single OSPF process.

Authentication

OSPF can authenticate OSPF packets. Only packets that pass the authentication are received. If an incoming hello packet cannot pass authentication, the neighbor relationship cannot be established.

The authentication type for interfaces attached to a single area must be identical. Authentication types include non-authentication, plaintext authentication, and MD5 ciphertext authentication. The authentication password for interfaces that are attached to a network segment must be identical.

OSPF Graceful Restart

Graceful Restart (GR) ensures the continuity of packet forwarding when a routing protocol restarts or an active/standby switchover occurs:

- **GR Restarter**—Graceful restarting router. It must have GR capability.
- **GR Helper**—A neighbor of the GR Restarter. It helps the GR Restarter to complete the GR process.

After an OSPF GR Restarter restarts, it must perform the following tasks.

- Obtain OSPF neighbor information.
- Obtain the LSDB.

Before restart, the GR Restarter negotiates GR capability with GR Helpers. During the restart of the GR Restarter, GR Helpers still advertise their adjacencies with the GR Restarter. After restart, the GR Restarter sends GR Helpers an OSPF GR signal so that the GR Helpers do not reset their neighbor relationships with the GR Restarter. Upon receiving responses from neighbors, the GR Restarter creates the neighbor relationships.

After that, the GR Restarter synchronizes the LSDB with GR-capable neighbors, updates its routing table and forwarding table, and removes stale routes.

BFD

Bidirectional forwarding detection (BFD) provides a single mechanism to quickly detect and monitor the connectivity of links between OSPF neighbors, reducing network convergence time. For more information about BFD, see *High Availability Configuration Guide*.

Protocols and standards

- RFC 1765, *OSPF Database Overflow*
- RFC 2328, *OSPF Version 2*
- RFC 3101, *OSPF Not-So-Stubby Area (NSSA) Option*
- RFC 3137, *OSPF Stub Router Advertisement*
- RFC 3630, *Traffic Engineering Extensions to OSPF Version 2*
- RFC 4811, *OSPF Out-of-Band LSDB Resynchronization*
- RFC 4812, *OSPF Restart Signaling*
- RFC 4813, *OSPF Link-Local Signaling*

OSPF configuration task list

Make a proper plan before configuring OSPF.

To run OSPF in a routing domain, you must first enable OSPF on the routers. Then use the default settings of parameters such as the hello interval, LSA delay timer, and SPF calculation interval. You can also configure them as needed. OSPF routers should be configured on an area basis. Wrong configurations may cause communication failures, routing information blocks, and routing loops.

Complete the following tasks to configure OSPF:

Task	Remarks
Enabling OSPF	Required
Configuring OSPF areas	Configuring a stub area
	Configuring an NSSA area
	Configuring a virtual link
Configuring OSPF network types	Configuring the OSPF network type for an interface as broadcast
	Configuring the OSPF network type for an interface as NBMA
	Configuring the OSPF network type for an interface as P2MP
	Configuring the OSPF network type for an interface as P2P
Configuring OSPF route control	Configuring OSPF route summarization
	Configuring OSPF inbound route filtering
	Configuring ABR Type-3 LSA filtering
	Configuring an OSPF cost for an interface
	Configuring the maximum number of OSPF routes
	Configuring the maximum number of ECMP routes

Task	Remarks	
Tuning and optimizing OSPF networks	Configuring OSPF preference	Optional
	Configuring OSPF route redistribution	Optional
	Configuring OSPF packet timers	Optional
	Specifying LSA transmission delay	Optional
	Specifying SPF calculation interval	Optional
	Specifying the LSA arrival interval	Optional
	Specifying the LSA generation interval	Optional
	Disabling interfaces from receiving and sending OSPF packets	Optional
	Configuring stub routers	Optional
	Configuring OSPF authentication	Optional
	Adding the interface MTU into DD packets	Optional
	Configuring the maximum number of external LSAs in LSDB	Optional
	Enabling compatibility with RFC 1583	Optional
	Logging neighbor state changes	Optional
	Configuring OSPF network management	Optional
	Enabling message logging	Optional
	Enabling the advertisement and reception of opaque LSAs	Optional
	Configuring OSPF to give priority to receiving and processing hello packets	Optional
	Configuring the LSU transmit rate	Optional
	Setting the DSCP value for OSPF packets	Optional
Enabling OSPF ISPF	Optional	
Configuring OSPF FRR	Optional	
Configuring OSPF Graceful Restart	Configuring the OSPF GR Restarter	Optional
	Configuring the OSPF GR Helper	Optional
	Triggering OSPF Graceful Restart	Optional
Configuring BFD for OSPF	Optional	

Enabling OSPF

You must enable OSPF before performing other OSPF configuration tasks.

Configuration prerequisites

Configure the link layer protocol, and IP addresses for interfaces so that neighboring nodes can reach each other.

Configuration procedure

To enable OSPF on a router, create an OSPF process and specify areas with which the process is associated, and the network segments contained in each area. If an interface's IP address resides on a network segment of an area, the interface belongs to the area and is enabled with OSPF, and OSPF advertises the direct route of the interface.

To run OSPF, a router must have a router ID, which is the unique identifier of the router in the AS. Following is additional information about router IDs:

- You can specify a router ID when creating the OSPF process. Any two routers in an AS must have different router IDs. In practice, the ID of a router is the IP address of one of its interfaces.
- If you specify no router ID when creating the OSPF process, the global router ID is used. HP recommends you to specify a router ID when you create the OSPF process.

OSPF can run multiple processes and supports VPNs as follows:

- When a router runs multiple OSPF processes, specify a router ID for each process, which takes effect locally and has no influence on packet exchange between routers. Two routers having different process IDs can exchange packets.
- OSPF support for VPNs enables an OSPF process to run in a specified VPN.

To enable OSPF:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure a global router ID.	router id <i>router-id</i>	Optional. Not configured by default. If no global router ID is configured, the highest loopback interface IP address, if any, is used as the router ID. If no loopback interface IP address is available, the highest physical interface IP address is used, regardless of the interface status.
3. Enable an OSPF process and enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	Not enabled by default.
4. Configure a description for the OSPF process.	description <i>description</i>	Optional. Not configured by default.
5. Configure an OSPF area and enter OSPF area view.	area <i>area-id</i>	Not configured by default.
6. Configure a description for the area.	description <i>description</i>	Optional. Not configured by default.
7. Specify a network to enable OSPF on the interface attached to the network.	network <i>ip-address</i> <i>wildcard-mask</i>	Not configured by default. A network segment can only belong to one area.

Configuring OSPF areas

After splitting an OSPF AS into multiple areas, configure some areas as stub areas or NSSA areas as needed.

If no connection can be achieved between the backbone and a non-backbone area, or within the backbone itself, you can configure virtual links to solve it.

Configuration prerequisites

Before you configure an OSPF area, complete the following tasks:

- Configure IP addresses for interfaces, and make sure that all neighboring nodes can reach each other at the network layer.
- Configure OSPF basic functions.

Configuring a stub area

Configure a non-backbone area at the AS edge as a stub area by configuring the **stub** command on all the routers attached to the area. In this way, Type-5 LSAs will not be flooded within the stub area, reducing the routing table size. The ABR generates a default route into the stub area so all packets destined outside of the AS are sent through the default route.

To further reduce the routing table size and routing information exchanged in the stub area, configure it as a totally stub area by using the **stub [no-summary]** command on the ABR. In this way, neither AS external routes nor inter-area routing information will be distributed into the area. All the packets destined outside of the AS or area will be sent to the ABR for forwarding.

To configure OSPF areas:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Enter area view.	area <i>area-id</i>	N/A
4. Configure the area as a stub area.	stub [default-route-advertise-always no-summary] *	Not configured by default. You must use the stub command on routers attached to a stub area. The backbone area cannot be a totally stub area. A (totally) stub area cannot have an ASBR because AS external routes cannot be distributed into the stub area.
5. Specify a cost for the default route advertised to the stub area.	default-cost <i>cost</i>	Optional. 1 by default. The default-cost command takes effect only on the ABR of a stub area.

NOTE:

Virtual links cannot transit totally stub areas.

Configuring an NSSA area

A stub area cannot redistribute routes. Configure the area as an NSSA area to allow for route redistribution by keeping other stub area characteristics.

To configure an NSSA area:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Enter area view.	area <i>area-id</i>	N/A
4. Configure the area as an NSSA area.	nssa [default-route-advertise no-import-route no-summary translate-always translator-stability-interval <i>value</i>] *	Not configured by default. You must use the nssa command on all the routers attached to an NSSA area.
5. Specify a cost for the default route advertised to the NSSA area.	default-cost <i>cost</i>	Optional. 1 by default. The default-cost command is available only on the ABR/ASBR of an NSSA area.

Configuring a virtual link

Non-backbone areas exchange routing information via the backbone area. Connectivity between the backbone and non-backbone areas and within the backbone must be maintained.

If the required physical links are not available for this connectivity maintenance, you can configure virtual links to resolve it.

To configure a virtual link:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Enter area view.	area <i>area-id</i>	N/A

Step	Command	Remarks
4. Configure a virtual link.	<code>vlink-peer router-id [hello seconds retransmit seconds trans-delay seconds dead seconds simple [plain cipher] password { md5 hmac-md5 } key-id [plain cipher] password] *</code>	You must configure this command on both ends of a virtual link. hello and dead intervals must be identical on both ends of the virtual link.

Configuring OSPF network types

OSPF classifies networks into the following types by the link layer protocol:

- **Broadcast**—When the link layer protocol is Ethernet or FDDI, OSPF considers the network type as broadcast by default.
- **NBMA**—When the link layer protocol is Frame Relay, ATM, or X.25, OSPF considers the network type as NBMA by default.
- **P2P**—When the link layer protocol is PPP, LAPB, or HDLC, OSPF considers the network type as P2P by default.

The following are examples of how you can change the network type of an interface as needed:

- When an NBMA network becomes fully meshed through address mapping—any two routers in the network have a direct virtual circuit in between, you can change the network type to broadcast to avoid manual configuration of neighbors.
- When some routers in the broadcast network do not support multicast, you can change the network type to NBMA.
- An NBMA network must be fully meshed. If it is partially meshed, you can change the network type to P2MP to simplify configuration and save network costs.
- If a router on an NBMA network has only one neighbor, you can change the network type to P2P to save network costs.

If two interfaces on a link are both configured as the broadcast, NBMA, or P2MP network type, they cannot establish a neighbor relationship unless they are on the same network segment.

Configuration prerequisites

Before you configure OSPF network types, complete the following tasks:

- Configure IP addresses for interfaces, making neighboring nodes accessible with each other at network layer.
- Configure OSPF basic functions.

Configuring the OSPF network type for an interface as broadcast

Step	Command	Remarks
1. Enter system view.	<code>system-view</code>	N/A
2. Enter interface view.	<code>interface interface-type interface-number</code>	N/A

Step	Command	Remarks
3.	Configure the OSPF network type for the interface as broadcast. ospf network-type broadcast	By default, the network type of an interface depends on the link layer protocol.
4.	Configure a router priority for the interface. ospf dr-priority priority	Optional. The default router priority is 1.

Configuring the OSPF network type for an interface as NBMA

After configuring the network type of an interface as NBMA, you must make some special configurations.

Because NBMA interfaces cannot find neighbors via broadcasting hello packets, you must specify neighbors and their router priorities. (A router priority of 0 means the router does not have the DR election right. A router priority greater than 0 means the router has the DR election right.)

The router priority configured with the **ospf dr-priority** command is for actual DR election. The priority configured with the **peer** command indicates whether a neighbor has the election right or not. If you configure the router priority for a neighbor as 0, the local router will assume the neighbor has no election right, and thus send no hello packets to this neighbor. However, if the local router is the DR or BDR, it still sends hello packets to the neighbor with priority 0 for neighborhood establishment.

To configure the OSPF network type for an Interface as NBMA:

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Enter interface view. interface interface-type interface-number	N/A
3.	Configure the OSPF network type for the interface as NBMA. ospf network-type nbma	By default, the network type of an interface depends on the link layer protocol.
4.	Configure a router priority for the interface. ospf dr-priority priority	Optional. The default router priority is 1.
5.	Exit to system view. quit	N/A
6.	Enter OSPF view. ospf [process-id router-id router-id vpn-instance vpn-instance-name] *	N/A
7.	Specify a neighbor and its router priority. peer ip-address [cost value dr-priority dr-priority]	N/A

Configuring the OSPF network type for an interface as P2MP

Step	Command	Remarks
1.	Enter system view. system-view	N/A

Step	Command	Remarks
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the OSPF network type for the interface as P2MP.	ospf network-type p2mp [unicast]	By default, the network type of an interface depends on the link layer protocol. After you configure the OSPF network type for an interface as P2MP unicast, all packets are unicast over the interface. The interface cannot broadcast hello packets to discover neighbors. In that case, you need to manually specify the neighbors.
4. Exit to system view.	quit	N/A
5. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
6. Specify a neighbor and its router priority on a P2MP unicast network.	peer <i>ip-address</i> [cost <i>value</i> dr-priority <i>dr-priority</i>]	Required if the interface type is P2MP unicast.

Configuring the OSPF network type for an interface as P2P

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the OSPF network type for the interface as P2P.	ospf network-type p2p	By default, the network type of an interface depends on the link layer protocol.

Configuring OSPF route control

This section describes how to control the advertisement and reception of OSPF routing information, as well as route redistribution from other protocols.

Configuration prerequisites

Before you configure OSPF route control, complete the following tasks:

- Configure IP addresses for interfaces.
- Configure OSPF basic functions.
- Configure filters if routing information filtering is needed.

Configuring OSPF route summarization

Route summarization is when an ABR or ASBR summarizes routes with the same prefix into a single route and distributes it to other areas.

Route summarization reduces the traffic of routing information exchanged between areas and the sizes of routing tables on routers, improving route calculation speed on routers.

For example, there are three internal routes in an area: 19.1.1.0/24, 19.1.2.0/24, and 19.1.3.0/24. By configuring route summarization on the ABR, the three routes are summarized into the route 19.1.0.0/16 that is advertised to other areas.

Configuring route summarization on an ABR

If contiguous network segments are available in the area, you can summarize them into a single network segment. An ABR generates Type-3 LSAs on a per network segment basis for an attached non-backbone area.

The ABR in the area distributes only the summary LSA to reduce the scale of LSDBs on routers in other areas and the influence of topology changes.

To configure route summarization on an ABR:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Enter OSPF area view.	area <i>area-id</i>	N/A
4. Configure ABR route summarization.	abr-summary <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } [advertise not-advertise] [cost <i>cost</i>]	Not configured by default. The command is available on an ABR only.

Configuring route summarization when redistributing routes into OSPF on an ASBR

Without route summarization, an ASBR advertises each redistributed route in a separate ASE LSA. After a summary route is configured, the ASBR advertises only the summary route in an ASE LSA instead of more specific routes, which reduces the number of LSAs in the LSDB.

The ASBR summarizes redistributed Type-5 LSAs that fall into the specified address range. If the ASBR is in an NSSA area, it also summarizes Type-7 LSAs that fall into the specified address range. If the ASBR is also the ABR, it summarizes Type-5 LSAs translated from Type-7 LSAs.

To configure route summarization when redistributing routes into OSPF on an ASBR:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Configure ASBR route summarization.	asbr-summary <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } [tag <i>tag</i> not-advertise cost <i>cost</i>] *	The command is available on an ASBR only. Not configured by default.

Configuring OSPF inbound route filtering

OSPF calculates routes by using LSAs. The calculated routes can be filtered and only permitted routes are installed into the OSPF routing table.

OSPF provides the following filtering methods:

- Filters routing information by destination address through ACLs and IP address prefixes
- Filters routing information by next hop through the filtering criteria configured with the **gateway** keyword
- Filters routing information by destination address through ACLs and IP address prefixes and by next hop through the filtering criteria configured with the **gateway** keyword
- Filters routing information by routing policy specified by the **route-policy** keyword

For more information about IP prefix list and routing policy, see "[Configuring routing policies.](#)"

To configure inbound route filtering:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Configure inbound route filtering.	filter-policy { <i>acl-number</i> [gateway <i>ip-prefix-name</i>] gateway <i>ip-prefix-name</i> ip-prefix <i>ip-prefix-name</i> [gateway <i>ip-prefix-name</i>] route-policy <i>route-policy-name</i> } import	Not configured by default.

Configuring ABR Type-3 LSA filtering

You can configure an ABR to filter Type-3 LSAs advertised to an area.

To configure Type-3 LSA filtering on an ABR:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Enter area view.	area <i>area-id</i>	N/A
4. Configure ABR Type-3 LSA filtering.	filter { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } { import export }	Not configured by default.

Configuring an OSPF cost for an interface

You can configure an OSPF cost for an interface by using either of the following methods:

- Configure the cost value in interface view.
- Configure a bandwidth reference value for the interface. OSPF computes the cost with this formula: Interface OSPF cost = Bandwidth reference value (100 Mbps)/Interface bandwidth (Mbps). If the

calculated cost is greater than 65535, the value of 65535 is used. If the calculated cost is less than 1, the value of 1 is used.

If the cost value is not configured for an interface, OSPF computes the interface cost automatically.

To configure an OSPF cost for an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure an OSPF cost for the interface.	ospf cost <i>value</i>	Optional. The default cost depends on the interface type: 1 for a VLAN interface; 0 for a loopback interface; computed according to the bandwidth for other interfaces.

To configure a bandwidth reference value:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Configure a bandwidth reference value.	bandwidth-reference <i>value</i>	Optional. The value defaults to 100 Mbps.

Configuring the maximum number of OSPF routes

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Configure the maximum number of OSPF routes.	maximum-routes { external inter intra } <i>number</i>	Optional. The number varies with devices.

Configuring the maximum number of ECMP routes

Perform this task to implement load sharing over ECMP routes.

To configure the maximum number of ECMP routes:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Configure the maximum number of ECMP routes.	maximum load-balancing <i>maximum</i>	Optional. 8 by default.

Configuring OSPF preference

A router can run multiple routing protocols, and each protocol is assigned a preference. When the routing protocols find routes to the same destination, the route found by the protocol with the highest preference is selected as the best route.

To configure OSPF preference:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Configure a preference for OSPF.	preference [ase] [route-policy <i>route-policy-name</i>] <i>value</i>	Optional. By default, the preference of OSPF internal routes is 10, and the preference of OSPF external routes is 150.

Configuring OSPF route redistribution

This section describes configuring OSPF to redistribute manually configured routes or routes discovered by other routing protocols.

Only active routes can be redistributed. Use the **display ip routing-table protocol** command to view route state information.

Configuring route redistribution into OSPF

On a router running OSPF and other routing protocols, you can configure OSPF to redistribute routes from other protocols such as RIP, IS-IS, BGP, static, and direct routes, and advertise them in Type-5 LSAs or Type-7 LSAs.

By filtering redistributed routes, OSPF translates only permitted routes into Type-5 LSAs or Type-7 LSAs for advertisement.

To configure OSPF route redistribution:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A

Step	Command	Remarks
3. Configure OSPF to redistribute routes from another protocol.	import-route <i>protocol</i> [<i>process-id</i> all-processes allow-ibgp] [cost <i>cost</i> type <i>type</i> tag <i>tag</i> route-policy <i>route-policy-name</i>] *	Not configured by default
4. Configure OSPF to filter redistributed routes before advertisement.	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } export [<i>protocol</i> [<i>process-id</i>]]	Optional Not configured by default

Configuring OSPF to redistribute a default route

Using the **import-route** command cannot redistribute a default external route. To do so, you must use the **default-route-advertise** command.

To configure OSPF to redistribute a default external route:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Redistribute a default route.	default-route-advertise [[[always permit-calculate-other] cost <i>cost</i> route-policy <i>route-policy-name</i> type <i>type</i>] * summary <i>cost</i>]	Not redistributed by default. The default-route-advertise summary cost command is applicable only to VPN, and the default route is redistributed in a Type-3 LSA. The PE router advertises the default route to the CE router.

Configuring the default parameters for redistributed routes

You can configure default parameters, such as the cost, upper limit, tag and type for redistributed routes. Tags indicate information related to protocols. For example, when redistributing BGP routes, OSPF uses tags to identify AS IDs.

To configure the default parameters for redistributed routes:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A

Step	Command	Remarks
3.	Configure the default parameters for redistributed routes (cost, route number, tag, and type). default { cost <i>cost</i> limit <i>limit</i> tag <i>tag</i> type <i>type</i> } *	Optional. The default cost is 1, the default maximum number of routes redistributed per time is 1000, the default tag is 1, and default type of redistributed routes is Type-2.

Advertising a host route

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Enter OSPF view. ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3.	Enter area view. area <i>area-id</i>	N/A
4.	Advertise a host route. host-advertise <i>ip-address cost</i>	Optional. Not advertised by default.

Tuning and optimizing OSPF networks

You can optimize your OSPF network in the following ways:

- Change OSPF packet timers to adjust the OSPF network convergence speed and network load. On low-speed links, consider the delay time for sending LSAs.
- Change the SPF calculation interval to reduce resource consumption caused by frequent network changes.
- Configure OSPF authentication to improve security.
- Configure OSPF network management functions, such as binding OSPF MIB with a process, sending trap information, and collecting log information.

Configuration prerequisites

Before you configure OSPF network optimization, complete the following tasks:

- Configure IP addresses for interfaces.
- Configure OSPF basic functions.

Configuring OSPF packet timers

You can configure the following timers on OSPF interfaces as needed.

- **Hello timer**—Interval for sending hello packets. It must be identical on OSPF neighbors. The longer the interval, the lower the convergence speed, and the smaller the network load.

- **Poll timer**—Interval for sending hello packets to a neighbor that is down on the NBMA network. The poll interval is at least four times the hello interval.
- **Dead timer**—Interval within which if the interface receives no hello packet from the neighbor, it declares the neighbor is down. The dead interval must be at least four times the hello interval on an interface.
- **LSA retransmission timer**—Interval within which if the interface receives no acknowledgement packets after sending an LSA to the neighbor, it retransmits the LSA. An interval setting that is too small can cause unnecessary LSA retransmissions. This interval is typically set bigger than the round-trip time of a packet between two neighbors.

To configure timers for OSPF packets:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Specify the hello interval.	ospf timer hello <i>seconds</i>	Optional. The hello interval defaults to 10 seconds on P2P and broadcast interfaces, and defaults to 30 seconds on P2MP and NBMA interfaces. The default hello interval is restored when the network type for an interface is changed.
4. Specify the poll interval.	ospf timer poll <i>seconds</i>	Optional. The poll interval defaults to 120 seconds.
5. Specify the dead interval.	ospf timer dead <i>seconds</i>	Optional. The default dead interval is 40 seconds on P2P and broadcast interfaces and 120 seconds on P2MP and NBMA interfaces. The default dead interval is restored when the network type for an interface is changed.
6. Specify the retransmission interval.	ospf timer retransmit <i>interval</i>	Optional. The retransmission interval defaults to 5 seconds.

Specifying LSA transmission delay

Each LSA in the LSDB has an age that is incremented by 1 every second, but the age does not change during transmission. It is necessary to add a transmission delay into the age time especially for low-speed links.

To specify the LSA transmission delay on an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A

Step	Command	Remarks
3. Specify the LSA transmission delay.	ospf trans-delay <i>seconds</i>	Optional. 1 second by default.

Specifying SPF calculation interval

LSDB changes lead to SPF calculations. When the topology changes frequently, a large amount of network and router resources are occupied by SPF calculation. Adjust the SPF calculation interval to reduce the impact.

When network changes are not frequent, the *minimum-interval* is adopted. If network changes become frequent, the SPF calculation interval is incremented by *incremental-interval* × 2ⁿ⁻² (n is the number of calculation times) each time a calculation occurs until the *maximum-interval* is reached.

To configure SPF calculation interval:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Specify the SPF calculation interval.	spf-schedule-interval <i>maximum-interval</i> [<i>minimum-interval</i> [<i>incremental-interval</i>]]	Optional. By default, the interval is 5 seconds.

Specifying the LSA arrival interval

After receiving the same LSA as the previously received LSA within the LSA arrival interval, OSPF discards the LSA.

To configure the LSA arrival interval:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Configure the LSA arrival interval.	lsa-arrival-interval <i>interval</i>	Optional. 1000 milliseconds by default. Make sure this interval is smaller than or equal to the interval set with the lsa-generation-interval command.

Specifying the LSA generation interval

You can adjust the LSA generation interval to protect network resources and routers from being over consumed by frequent network changes.

When network changes are not frequent, LSAs are generated at the *minimum-interval*. If network changes become frequent, the LSA generation interval is incremented by *incremental-interval* × 2ⁿ⁻² (n is the number of generation times) each time a LSA generation occurs until the *maximum-interval* is reached.

To configure the LSA generation interval:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Configure the LSA generation interval.	lsa-generation-interval <i>maximum-interval</i> [<i>initial-interval</i> [<i>incremental-interval</i>]]	Optional. By default, the maximum interval is 5 seconds, the minimum interval is 0 milliseconds, and the incremental interval is 5000 milliseconds.

Disabling interfaces from receiving and sending OSPF packets

Follow these guidelines when you disable interfaces from receiving and sending OSPF packets:

- Different OSPF processes can disable the same interface from receiving and sending OSPF packets. The **silent-interface** command disables only the interfaces associated with the current process rather than interfaces associated with other processes.
- After an OSPF interface is set to **silent**, other interfaces on the router can advertise direct routes of the interface in Router LSAs, but the interface cannot send any packet. This configuration can enhance OSPF adaptability and reduce resource consumption.

To disable interfaces from receiving and sending routing information:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Disable interfaces from receiving and sending OSPF packets.	silent-interface { <i>interface-type</i> <i>interface-number</i> all }	Optional. Not disabled by default.

Configuring stub routers

A stub router is used for traffic control. It tells other OSPF routers to not use it to forward data.

The Router LSAs from the stub router may contain different link type values. A value of 3 means a link to a stub network, and the cost of the link will not be changed. A value of 1, 2, or 4 means a point-to-point link, a link to a transit network, or a virtual link. On such links, a maximum cost value of 65535 is used. Thus, neighbors find that the links to the stub router have large costs that they will not send packets to the stub router for forwarding as long as another route with a smaller cost exists.

To configure a router as a stub router:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Configure the router as a stub router.	stub-router	By default, the router is not a stub router in any OSPF process.

Configuring OSPF authentication

You can configure OSPF packet authentication to ensure the security of packet exchanges.

After authentication is configured, OSPF only receives packets that pass authentication. Failed packets cannot establish neighboring relationships.

To configure OSPF authentication, you must configure the same area authentication mode on all the routers in the area. In addition, the authentication mode and password for all interfaces attached to the same area must be identical.

To configure OSPF authentication:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Enter area view.	area <i>area-id</i>	N/A
4. Configure the authentication mode.	authentication-mode { md5 simple }	Not configured by default.
5. Return to OSPF view.	quit	N/A
6. Return to system view.	quit	N/A
7. Enter interface view.	interface <i>interface-type interface-number</i>	N/A
8. Configure interface authentication mode.	<ul style="list-style-type: none"> Configure the simple authentication mode: ospf authentication-mode simple [cipher plain] <i>password</i> Configure the MD5 authentication mode: ospf authentication-mode { hmac-md5 md5 } <i>key-id</i> [cipher plain] <i>password</i> 	Use either approach. Not configured by default.

Adding the interface MTU into DD packets

By default, an interface adds 0 into the interface MTU field of a DD packet to be sent rather than the interface MTU. You can enable an interface to add its MTU into DD packets.

To add the interface MTU into DD packets:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable the interface to add its MTU into DD packets.	ospf mtu-enable	Optional. Not enabled by default.

Configuring the maximum number of external LSAs in LSDB

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Specify the maximum number of external LSAs in the LSDB.	lsdb-overflow-limit <i>number</i>	Optional. Not specified by default.

Enabling compatibility with RFC 1583

RFC 1583 specifies a different method than RFC 2328 for selecting an external route from multiple LSAs. If RFC 1583 is made compatible with RFC 2328, the routes in the backbone area are preferred; if not, the routes in the non-backbone area are preferred to reduce the burden of the backbone area.

To avoid routing loops, HP recommends configuring all the routers to be either compatible or incompatible with RFC 1583.

To make them compatible:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Enable compatibility with RFC 1583.	rfc1583 compatible	Optional. Enabled by default.

Logging neighbor state changes

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Enable the logging of neighbor state changes.	log-peer-change	Optional. Enabled by default.

Configuring OSPF network management

With trap generation enabled, OSPF generates traps to report important events. Traps fall into the following levels.

- **Level-3**—Fault traps
- **Level-4**—Alarm traps
- **Level-5**—Normal but important traps
- **Level-6**—Notification traps

The generated traps are sent to the information center of the device. The output rules of the traps such as whether to output the traps and the output direction are determined according to the information center configuration. (For information center configuration, see *Network Management and Monitoring Configuration Guide*.)

To configure OSPF network management:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Bind OSPF MIB to an OSPF process.	ospf mib-binding <i>process-id</i>	Optional. By default, the OSPF process with the smallest <i>process-id</i> is bound with OSPF MIB.
3. Enable OSPF trap generation.	snmp-agent trap enable ospf [<i>process-id</i>] [ifaauthfail ifcfgerror ifrxbadpkt ifstatechange iftxretransmit lsdbapproachoverflow lsdboverflow maxagelsa nbrstatechange originatelsa vifcfgerror virifaauthfail virifrxbadpkt virifstatechange viriftxretransmit virnbrstatechange] *	Optional. Enabled by default.

Enabling message logging

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Enable message logging.	enable log [config error state]	Not enabled by default.

Enabling the advertisement and reception of opaque LSAs

With this feature enabled, OSPF can receive and advertise Type 9, Type 10, and Type 11 opaque LSAs.

To enable the advertisement and reception of opaque LSAs:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Enable the advertisement and reception of opaque LSAs.	opaque-capability enable	Optional. Disabled by default.

Configuring OSPF to give priority to receiving and processing hello packets

To ensure OSPF runs properly, a router receives and processes hello packets and other protocol packets at the same time. When the router has established neighbor relationships with multiple routers, and the routing table size is big, the router must receive and process large numbers of packets. In this case, you can configure OSPF to give priority to receiving and processing hello packets to ensure stable neighbor relationships.

To configure OSPF to give priority to receiving and processing hello packets:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure OSPF to give priority to receiving and processing hello packets.	ospf packet-process prioritized-treatment	Not configured by default.

Configuring the LSU transmit rate

Sending large numbers of LSU packets affects router performance and consumes too much network bandwidth. You can configure the router to send LSU packets at a proper interval and limit the maximum number of LSU packets sent out of an OSPF interface each time.

To configure the LSU transmit rate:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Configure the LSU transmit rate.	transmit-pacing interval <i>interval</i> count <i>count</i>	Optional. By default, an OSPF interface sends up to three LSU packets every 20 milliseconds.

Setting the DSCP value for OSPF packets

An IPv4 packet header contains an 8-bit TOS field. As defined in RFC 2474, the first six bits set the Differentiated Services Code Point (DSCP) value and the last two bits are reserved. Network devices use the DSCP value as a reference to determine the packet priority for transmission.

You can set the DSCP value for OSPF packets.

To set the DSCP value for OSPF packets:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable an OSPF process and enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	By default, OSPF is not enabled.
3. Set the DSCP value for OSPF packets.	dscp <i>dscp-value</i>	Optional. By default, the DSCP value in OSPF packets is 48.

Enabling OSPF ISPF

When a network topology is changed, Incremental Shortest Path First (ISPF) allows the system to recompute only the affected part of the shortest path tree (SPT), instead of the entire SPT.

To enable OSPF ISPF:

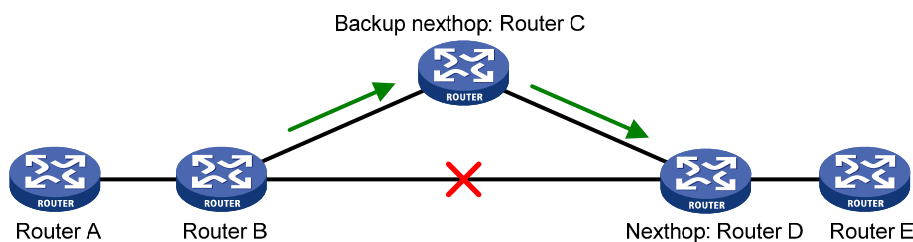
Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Enable OSPF ISPF.	ispf enable	Optional. Disabled by default.

Configuring OSPF FRR

When a link fails, the packets on the link are discarded, and a routing loop may occur until OSPF completes routing convergence based on the new network topology.

You can enable OSPF fast reroute (FRR) to solve this problem.

Figure 37 Network diagram for OSPF FRR



In [Figure 37](#), after you enable FRR on Router B, OSPF automatically calculates or designates a backup next hop when a link failure is detected. Packets are directed to the backup next hop. At the same time, OSPF calculates the shortest path based on the new network topology, and forwards packets over the path after routing convergence.

You can either configure OSPF FRR to calculate a backup next hop automatically, or to designate a backup next hop by using a routing policy.

Configuration prerequisites

Before you configure OSPF FRR, complete the following tasks:

- Configure network layer addresses for interfaces.
- Enable OSPF.

Configuration guidelines

- Do not use FRR and BFD at the same time. Otherwise, FRR may fail to take effect.
- Do not use the **fast-reroute auto** command together with the commands **vlink-peer**.

Configuring OSPF FRR to automatically calculate a backup next hop

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the source address of echo packets.	bfd echo-source-ip <i>ip-address</i>	Not configured by default.
3. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
4. Enable OSPF FRR to automatically calculate a backup next hop.	fast-reroute auto [abr-only]	Not configured by default. If abr-only is specified, the route to the ABR is selected as the backup path.

Configuring OSPF FRR to designate a backup next hop with a routing policy

Before configuring this task, you need to use the **apply fast-reroute backup-interface** command to specify a backup next hop in the routing policy to be referenced. For more information about the **apply fast-reroute backup-interface** command and routing policy configuration, see "[Configuring routing policies](#)."

To configure OSPF FRR:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the source address of echo packets.	bfd echo-source-ip <i>ip-address</i>	Not configured by default.
3. Enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
4. Enable OSPF FRR to designate a backup next hop by using a routing policy.	fast-reroute route-policy <i>route-policy-name</i>	Not configured by default.

Configuring OSPF Graceful Restart

OSPF GR involves the following:

- **IETF standard GR**—Uses Opaque LSAs to implement GR.
- **Non IETF standard GR**—Uses link local signaling (LLS) to advertise GR capability and uses out of band synchronization to synchronize the LSDB.

A device can act as a GR Restarter and GR Helper at the same time.

Configuring the OSPF GR Restarter

You can configure the IETF standard or non IETF standard OSPF GR Restarter.

Configuring the IETF standard OSPF GR Restarter

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable OSPF and enter its view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Enable opaque LSA advertisement capability.	opaque-capability enable	Disabled by default.
4. Enable the IETF standard Graceful Restart capability.	graceful-restart ietf	Disabled by default.
5. Configure the Graceful Restart interval.	graceful-restart interval <i>timer</i>	Optional. 120 seconds by default.

Configuring the non-IETF standard OSPF GR Restarter

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable OSPF and enter its view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Enable the link-local signaling capability.	enable link-local-signaling	Disabled by default.
4. Enable the out-of-band re-synchronization capability.	enable out-of-band-resynchronization	Disabled by default.
5. Enable non-IETF standard Graceful Restart capability for OSPF.	graceful-restart [nonstandard]	Disabled by default.
6. Configure Graceful Restart interval for OSPF.	graceful-restart interval <i>timer</i>	Optional. 120 seconds by default.

Configuring the OSPF GR Helper

You can configure the IETF standard or non-IETF standard OSPF GR Helper.

Configuring the IETF standard OSPF GR Helper

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable OSPF and enter its view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Enable opaque LSA reception and advertisement.	opaque-capability enable	Not enabled by default.
4. Configure the neighbors for which the router can serve as a GR Helper.	graceful-restart help { <i>acl-number</i> prefix <i>prefix-list</i> }	Optional. The router can serve as a GR Helper for any OSPF neighbor by default.

Configuring the non IETF standard OSPF GR Helper

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable OSPF and enter its view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A
3. Enable the link-local signaling capability.	enable link-local-signaling	Disabled by default.
4. Enable the out-of-band re-synchronization capability.	enable out-of-band-resynchronization	Disabled by default.
5. Configure the neighbors for which the router can serve as a GR Helper.	graceful-restart help { <i>acl-number</i> prefix <i>prefix-list</i> }	Optional. The router can serve as a GR Helper for any OSPF neighbor by default.

Triggering OSPF Graceful Restart

Performing a Master/Slave switchover, or performing the following configuration on an OSPF router, will trigger OSPF Graceful Restart.

To trigger OSPF Graceful Restart:

Task	Command	Remarks
Trigger OSPF Graceful Restart.	reset ospf [<i>process-id</i>] process graceful-restart	Available in user view

Configuring BFD for OSPF

OSPF supports the following BFD detection methods:

- Control packet bidirectional detection, which requires BFD configuration to be made on both OSPF routers on the link.
- Echo packet single-hop detection, which requires BFD configuration to be made on one OSPF router on the link.

Configuring control packet bidirectional detection

Both ends of a BFD session must be on the same network segment and in the same area. One network segment can only belong to one area.

To enable BFD control packet bidirectional detection on an OSPF interface:

Step	Command	Description
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable BFD control packet bidirectional detection on the interface.	ospf bfd enable	Not enabled by default.

Configuring echo packet single-hop detection

Step	Command	Description
1. Enter system view.	system-view	N/A
2. Configure the source address of echo packets.	bfd echo-source-ip <i>ip-address</i>	Not configured by default.
3. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Enable BFD echo packet single-hop detection on the interface.	ospf bfd enable echo	Not enabled by default.

Displaying and maintaining OSPF

Task	Command	Remarks
Display OSPF brief information.	display ospf [<i>process-id</i>] brief [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPF statistics.	display ospf [<i>process-id</i>] cumulative [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display Link State Database information.	display ospf [<i>process-id</i>] lsdb [brief [{ ase router network summary asbr nssa opaque-link opaque-area opaque-as } [<i>link-state-id</i>]] [originate-router advertising-router-id self-originate]] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPF neighbor information.	display ospf [<i>process-id</i>] peer [verbose] [<i>interface-type interface-number</i>] [<i>neighbor-id</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display neighbor statistics of OSPF areas.	display ospf [<i>process-id</i>] peer statistics [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display next hop information.	display ospf [<i>process-id</i>] nexthop [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display routing table information.	display ospf [<i>process-id</i>] routing [interface <i>interface-type interface-number</i>] [nexthop <i>nexthop-address</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display virtual link information.	display ospf [<i>process-id</i>] vlink [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPF request queue information.	display ospf [<i>process-id</i>] request-queue [<i>interface-type interface-number</i>] [<i>neighbor-id</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPF retransmission queue information.	display ospf [<i>process-id</i>] retrans-queue [<i>interface-type interface-number</i>] [<i>neighbor-id</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPF ABR and ASBR information.	display ospf [<i>process-id</i>] abr-asbr [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPF interface information.	display ospf [<i>process-id</i>] interface [all <i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPF error information.	display ospf [<i>process-id</i>] error [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPF ASBR summarization information.	display ospf [<i>process-id</i>] asbr-summary [<i>ip-address</i> { <i>mask</i> <i>mask-length</i> }] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the global router ID.	display router id [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Reset OSPF counters.	reset ospf [<i>process-id</i>] counters [neighbor <i>interface-type interface-number</i>] [<i>router-id</i>]]	Available in user view
Reset an OSPF process.	reset ospf [<i>process-id</i>] process [graceful-restart]	Available in user view
Re-enable OSPF route redistribution.	reset ospf [<i>process-id</i>] redistribution	Available in user view

OSPF configuration examples

These examples only cover commands for OSPF configuration.

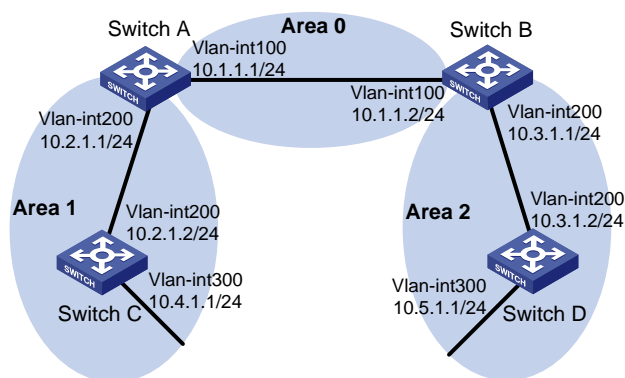
Configuring OSPF basic functions

Network requirements

As shown in [Figure 38](#), all switches run OSPF. The AS is split into three areas, where Switch A and Switch B act as ABRs to forward routing information between areas.

After configuration, all switches can learn routes to every network segment in the AS.

Figure 38 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure OSPF basic functions:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] ospf
[SwitchA-ospf-1] area 0
[SwitchA-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] quit
[SwitchA-ospf-1] area 1
[SwitchA-ospf-1-area-0.0.0.1] network 10.2.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.1] quit
[SwitchA-ospf-1] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] ospf
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] area 2
[SwitchB-ospf-1-area-0.0.0.2] network 10.3.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.2] quit
[SwitchB-ospf-1] quit
```

Configure Switch C

```
<SwitchC> system-view
[SwitchC] ospf
[SwitchC-ospf-1] area 1
[SwitchC-ospf-1-area-0.0.0.1] network 10.2.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.1] network 10.4.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.1] quit
[SwitchC-ospf-1] quit
```

Configure Switch D

```
<SwitchD> system-view
[SwitchD] ospf
[SwitchD-ospf-1] area 2
[SwitchD-ospf-1-area-0.0.0.2] network 10.3.1.0 0.0.0.255
[SwitchD-ospf-1-area-0.0.0.2] network 10.5.1.0 0.0.0.255
[SwitchD-ospf-1-area-0.0.0.2] quit
[SwitchD-ospf-1] quit
```

3. Verify the configuration:

Display information about neighbors on Switch A.

```
[SwitchA] display ospf peer verbose
```

```
OSPF Process 1 with Router ID 10.2.1.1
```

```
Neighbors
```

```
Area 0.0.0.0 interface 10.1.1.1(Vlan-interface100)'s neighbors
```

```
Router ID: 10.3.1.1          Address: 10.1.1.2          GR State: Normal
```

```
State: Full Mode: Nbr is Master Priority: 1
```

```
DR: 10.1.1.1 BDR: 10.1.1.2 MTU: 0
```

```
Dead timer due in 37 sec
```

```
Neighbor is up for 06:03:59
```

```
Authentication Sequence: [ 0 ]
```

```
Neighbor state change count: 5
```

```
Neighbors
```

```
Area 0.0.0.1 interface 10.2.1.1(Vlan-interface200)'s neighbors
```

```
Router ID: 10.4.1.1          Address: 10.2.1.2          GR State: Normal
```

```
State: Full Mode: Nbr is Master Priority: 1
```

```
DR: 10.2.1.1 BDR: 10.2.1.2 MTU: 0
```

```
Dead timer due in 32 sec
```

```
Neighbor is up for 06:03:12
```

```
Authentication Sequence: [ 0 ]
```

```
Neighbor state change count: 5
```

Display OSPF routing information on Switch A.

```
[SwitchA] display ospf routing
```

```
OSPF Process 1 with Router ID 10.2.1.1
```

```
Routing Tables
```

Routing for Network

Destination	Cost	Type	NextHop	AdvRouter	Area
10.2.1.0/24	10	Transit	10.2.1.1	10.2.1.1	0.0.0.1
10.3.1.0/24	4	Inter	10.1.1.2	10.3.1.1	0.0.0.0
10.4.1.0/24	13	Stub	10.2.1.2	10.4.1.1	0.0.0.1
10.5.1.0/24	14	Inter	10.1.1.2	10.3.1.1	0.0.0.0
10.1.1.0/24	2	Transit	10.1.1.1	10.2.1.1	0.0.0.0

Total Nets: 5

Intra Area: 3 Inter Area: 2 ASE: 0 NSSA: 0

Display the Link State Database on Switch A.

[SwitchA] display ospf lsdb

OSPF Process 1 with Router ID 10.2.1.1

Link State Database

Area: 0.0.0.0

Type	LinkState ID	AdvRouter	Age	Len	Sequence	Metric
Router	10.2.1.1	10.2.1.1	1069	36	80000012	0
Router	10.3.1.1	10.3.1.1	780	36	80000011	0
Network	10.1.1.1	10.2.1.1	1069	32	80000010	0
Sum-Net	10.5.1.0	10.3.1.1	780	28	80000003	12
Sum-Net	10.2.1.0	10.2.1.1	1069	28	8000000F	10
Sum-Net	10.3.1.0	10.3.1.1	780	28	80000014	2
Sum-Net	10.4.1.0	10.2.1.1	769	28	8000000F	13

Area: 0.0.0.1

Type	LinkState ID	AdvRouter	Age	Len	Sequence	Metric
Router	10.2.1.1	10.2.1.1	769	36	80000012	0
Router	10.4.1.1	10.4.1.1	1663	48	80000012	0
Network	10.2.1.1	10.2.1.1	769	32	80000010	0
Sum-Net	10.5.1.0	10.2.1.1	769	28	80000003	14
Sum-Net	10.3.1.0	10.2.1.1	1069	28	8000000F	4
Sum-Net	10.1.1.0	10.2.1.1	1069	28	8000000F	2
Sum-Asbr	10.3.1.1	10.2.1.1	1069	28	8000000F	2

Display OSPF routing information on Switch D.

[SwitchD] display ospf routing

OSPF Process 1 with Router ID 10.5.1.1

Routing Tables

Routing for Network

Destination	Cost	Type	NextHop	AdvRouter	Area
10.2.1.0/24	22	Inter	10.3.1.1	10.3.1.1	0.0.0.2
10.3.1.0/24	10	Transit	10.3.1.2	10.3.1.1	0.0.0.2
10.4.1.0/24	25	Inter	10.3.1.1	10.3.1.1	0.0.0.2
10.5.1.0/24	10	Stub	10.5.1.1	10.5.1.1	0.0.0.2
10.1.1.0/24	12	Inter	10.3.1.1	10.3.1.1	0.0.0.2

```

Total Nets: 5
Intra Area: 2  Inter Area: 3  ASE: 0  NSSA: 0
# On Switch D, ping the IP address 10.4.1.1 to check connectivity.
[SwitchD] ping 10.4.1.1
  PING 10.4.1.1: 56 data bytes, press CTRL_C to break
    Reply from 10.4.1.1: bytes=56 Sequence=2 ttl=253 time=2 ms
    Reply from 10.4.1.1: bytes=56 Sequence=2 ttl=253 time=1 ms
    Reply from 10.4.1.1: bytes=56 Sequence=3 ttl=253 time=1 ms
    Reply from 10.4.1.1: bytes=56 Sequence=4 ttl=253 time=1 ms
    Reply from 10.4.1.1: bytes=56 Sequence=5 ttl=253 time=1 ms

--- 10.4.1.1 ping statistics ---
  5 packet(s) transmitted
  5 packet(s) received
  0.00% packet loss
  round-trip min/avg/max = 1/1/2 ms

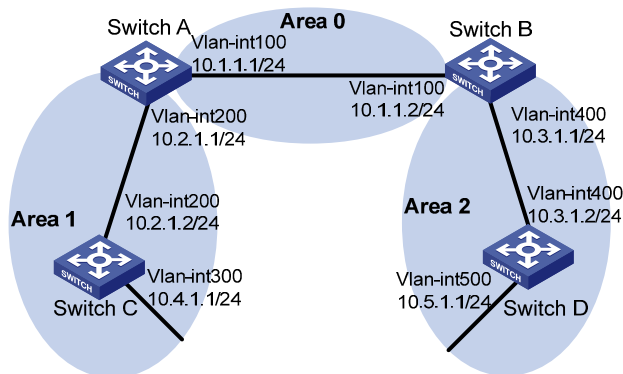
```

Configuring OSPF route redistribution

Network requirements

As shown in [Figure 39](#), all the switches run OSPF, and the AS is divided into three areas. Switch A and Switch B act as ABRs to forward routes between areas. Switch C is configured as an ASBR to redistribute external routes (static routes). Routing information is propagated properly in the AS.

Figure 39 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure OSPF basic functions. (See "[Configuring OSPF basic functions](#)")
3. Configure OSPF to redistribute routes:

```
# On Switch C, configure a static route destined for network 3.1.2.0/24.
```

```
<SwitchC> system-view
[SwitchC] ip route-static 3.1.2.1 24 10.4.1.2
```

```
# On Switch C, configure OSPF to redistribute static routes.
```

```
[SwitchC] ospf 1
```

```
[SwitchC-ospf-1] import-route static
```

4. Verify the configuration:

```
# Display the ABR/ASBR information of Switch D.
```

```
<SwitchD> display ospf abr-asbr
```

```
OSPF Process 1 with Router ID 10.5.1.1  
Routing Table to ABR and ASBR
```

Type	Destination	Area	Cost	NextHop	RtType
Intra	10.3.1.1	0.0.0.2	10	10.3.1.1	ABR
Inter	10.4.1.1	0.0.0.2	22	10.3.1.1	ASBR

```
# Display the OSPF routing table of Switch D.
```

```
<SwitchD> display ospf routing
```

```
OSPF Process 1 with Router ID 10.5.1.1  
Routing Tables
```

```
Routing for Network
```

Destination	Cost	Type	NextHop	AdvRouter	Area
10.2.1.0/24	22	Inter	10.3.1.1	10.3.1.1	0.0.0.2
10.3.1.0/24	10	Transit	10.3.1.2	10.3.1.1	0.0.0.2
10.4.1.0/24	25	Inter	10.3.1.1	10.3.1.1	0.0.0.2
10.5.1.0/24	10	Stub	10.5.1.1	10.5.1.1	0.0.0.2
10.1.1.0/24	12	Inter	10.3.1.1	10.3.1.1	0.0.0.2

```
Routing for ASEs
```

Destination	Cost	Type	Tag	NextHop	AdvRouter
3.1.2.0/24	1	Type2	1	10.3.1.1	10.4.1.1

```
Total Nets: 6
```

```
Intra Area: 2 Inter Area: 3 ASE: 1 NSSA: 0
```

Configuring OSPF to advertise a summary route

Network requirements

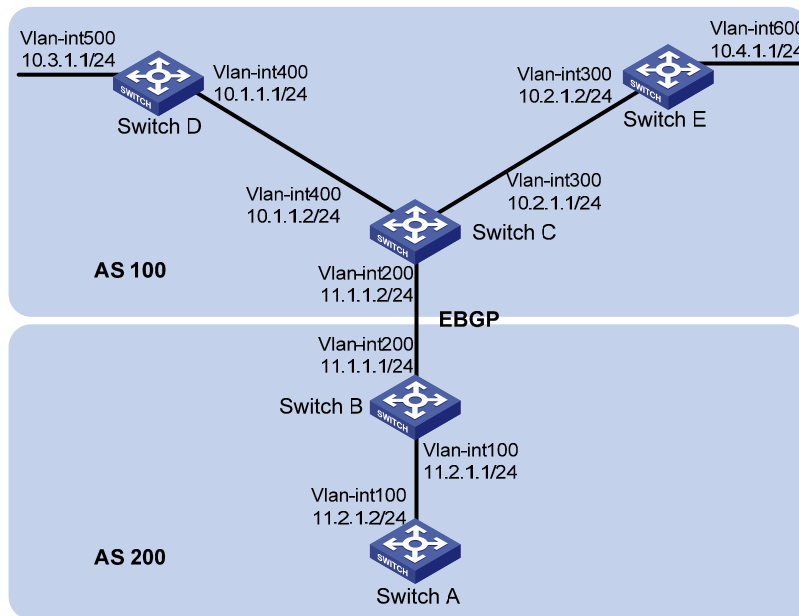
As shown in [Figure 40](#):

- Switch A and Switch B are in AS 200, which runs OSPF.
- Switch C, Switch D, and Switch E are in AS 100, which runs OSPF.

An EBGP connection is established between Switch B and Switch C. Switch B and Switch C are configured to redistribute OSPF routes and direct routes into BGP and BGP routes into OSPF.

Switch B is configured with route summarization and advertises only the summary route 10.0.0.0/8 to reduce Switch A's routing table size.

Figure 40 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure OSPF basic functions:
 - # Configure Switch A.

```
<SwitchA> system-view
[SwitchA] ospf
[SwitchA-ospf-1] area 0
[SwitchA-ospf-1-area-0.0.0.0] network 11.2.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] quit
[SwitchA-ospf-1] quit
```

 - # Configure Switch B.

```
<SwitchB> system-view
[SwitchB] ospf
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 11.2.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] quit
```

 - # Configure Switch C.

```
<SwitchC> system-view
[SwitchC] ospf
[SwitchC-ospf-1] area 0
[SwitchC-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] network 10.2.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] quit
[SwitchC-ospf-1] quit
```

 - # Configure Switch D.

```
<SwitchD> system-view
[SwitchD] ospf
```

```
[SwitchD-ospf-1] area 0
[SwitchD-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[SwitchD-ospf-1-area-0.0.0.0] network 10.3.1.0 0.0.0.255
[SwitchD-ospf-1-area-0.0.0.0] quit
```

Configure Switch E.

```
<SwitchE> system-view
[SwitchE] ospf
[SwitchE-ospf-1] area 0
[SwitchE-ospf-1-area-0.0.0.0] network 10.2.1.0 0.0.0.255
[SwitchE-ospf-1-area-0.0.0.0] network 10.4.1.0 0.0.0.255
[SwitchE-ospf-1-area-0.0.0.0] quit
[SwitchE-ospf-1] quit
```

3. Configure BGP to redistribute OSPF routes and direct routes:

Configure Switch B.

```
[SwitchB] bgp 200
[SwitchB-bgp] peer 11.1.1.2 as 100
[SwitchB-bgp] import-route ospf
[SwitchB-bgp] import-route direct
[SwitchB-bgp] quit
```

Configure Switch C.

```
[SwitchC] bgp 100
[SwitchC-bgp] peer 11.1.1.1 as 200
[SwitchC-bgp] import-route ospf
[SwitchC-bgp] import-route direct
[SwitchC-bgp] quit
```

4. Configure Switch B and Switch C to redistribute BGP routes into OSPF:

Configure OSPF to redistribute routes from BGP on Switch B.

```
[SwitchB] ospf
[SwitchB-ospf-1] import-route bgp
```

Configure OSPF to redistribute routes from BGP on Switch C.

```
[SwitchC] ospf
[SwitchC-ospf-1] import-route bgp
```

Display the OSPF routing table of Switch A.

```
[SwitchA] display ip routing-table
Routing Tables: Public
                Destinations : 8                Routes : 8
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
10.1.1.0/24	O_ASE	150	1	11.2.1.1	Vlan100
10.2.1.0/24	O_ASE	150	1	11.2.1.1	Vlan100
10.3.1.0/24	O_ASE	150	1	11.2.1.1	Vlan100
10.4.1.0/24	O_ASE	150	1	11.2.1.1	Vlan100
11.2.1.0/24	Direct	0	0	11.2.1.2	Vlan100
11.2.1.2/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0

- Configure summary route 10.0.0.0/8 on Switch B and advertise it:

```
[SwitchB-ospf-1] asbr-summary 10.0.0.0 8
```

Display the OSPF routing table of Switch A.

```
[SwitchA] display ip routing-table
```

```
Routing Tables: Public
```

```
Destinations : 5          Routes : 5
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
10.0.0.0/8	O_ASE	150	2	11.2.1.1	Vlan100
11.2.1.0/24	Direct	0	0	11.2.1.2	Vlan100
11.2.1.2/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0

The output shows that routes 10.1.1.0/24, 10.2.1.0/24, 10.3.1.0/24 and 10.4.1.0/24 are summarized into one route 10.0.0.0/8.

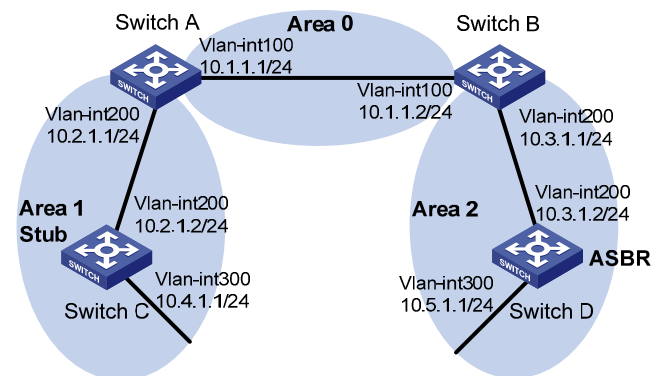
Configuring an OSPF stub area

Network requirements

Figure 41 shows an AS is split into three areas, where all switches run OSPF. Switch A and Switch B act as ABRs to forward routing information between areas. Switch D acts as the ASBR to redistribute routes (static routes).

Configure Area 1 as a stub area, reducing LSAs to this area without affecting route reachability.

Figure 41 Network diagram



Configuration procedure

- Configure IP addresses for interfaces. (Details not shown.)
- Configure OSPF basic functions. (See "Configuring OSPF basic functions")
- Configure Switch D to redistribute static routes:

```
<SwitchD> system-view
```

```
[SwitchD] ip route-static 3.1.2.1 24 10.5.1.2
```

```
[SwitchD] ospf
```

```
[SwitchD-ospf-1] import-route static
```

```
[SwitchD-ospf-1] quit
```

Display ABR/ASBR information on Switch C.

```
<SwitchC> display ospf abr-asbr
```

```
OSPF Process 1 with Router ID 10.4.1.1
Routing Table to ABR and ASBR
```

Type	Destination	Area	Cost	NextHop	RtType
Intra	10.2.1.1	0.0.0.1	3	10.2.1.1	ABR
Inter	10.3.1.1	0.0.0.1	5	10.2.1.1	ABR
Inter	10.5.1.1	0.0.0.1	7	10.2.1.1	ASBR

Display OSPF routing table information on Switch C.

```
<SwitchC> display ospf routing
```

```
OSPF Process 1 with Router ID 10.4.1.1
Routing Tables
```

Routing for Network

Destination	Cost	Type	NextHop	AdvRouter	Area
10.2.1.0/24	3	Transit	10.2.1.2	10.2.1.1	0.0.0.1
10.3.1.0/24	7	Inter	10.2.1.1	10.2.1.1	0.0.0.1
10.4.1.0/24	3	Stub	10.4.1.1	10.4.1.1	0.0.0.1
10.5.1.0/24	17	Inter	10.2.1.1	10.2.1.1	0.0.0.1
10.1.1.0/24	5	Inter	10.2.1.1	10.2.1.1	0.0.0.1

Routing for ASEs

Destination	Cost	Type	Tag	NextHop	AdvRouter
3.1.2.0/24	1	Type2	1	10.2.1.1	10.5.1.1

Total Nets: 6

Intra Area: 2 Inter Area: 3 ASE: 1 NSSA: 0

The output shows that because Switch C resides in a normal OSPF area, its routing table contains an external route.

4. Configure Area 1 as a Stub area:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] ospf
[SwitchA-ospf-1] area 1
[SwitchA-ospf-1-area-0.0.0.1] stub
[SwitchA-ospf-1-area-0.0.0.1] quit
[SwitchA-ospf-1] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] ospf
[SwitchC-ospf-1] area 1
[SwitchC-ospf-1-area-0.0.0.1] stub
[SwitchC-ospf-1-area-0.0.0.1] quit
[SwitchC-ospf-1] quit
```

```
# Display OSPF routing information on Switch C
```

```
[SwitchC] display ospf routing
```

```
OSPF Process 1 with Router ID 10.4.1.1
Routing Tables
```

```
Routing for Network
```

Destination	Cost	Type	NextHop	AdvRouter	Area
0.0.0.0/0	4	Inter	10.2.1.1	10.2.1.1	0.0.0.1
10.2.1.0/24	3	Transit	10.2.1.2	10.2.1.1	0.0.0.1
10.3.1.0/24	7	Inter	10.2.1.1	10.2.1.1	0.0.0.1
10.4.1.0/24	3	Stub	10.4.1.1	10.4.1.1	0.0.0.1
10.5.1.0/24	17	Inter	10.2.1.1	10.2.1.1	0.0.0.1
10.1.1.0/24	5	Inter	10.2.1.1	10.2.1.1	0.0.0.1

```
Total Nets: 6
```

```
Intra Area: 2 Inter Area: 4 ASE: 0 NSSA: 0
```

The output shows that when Switch C resides in the stub area, a default route takes the place of the external route.

```
# Filter Type-3 LSAs out the stub area
```

```
[SwitchA] ospf
```

```
[SwitchA-ospf-1] area 1
```

```
[SwitchA-ospf-1-area-0.0.0.1] stub no-summary
```

```
[SwitchA-ospf-1-area-0.0.0.1] quit
```

```
# Display OSPF routing information on Switch C.
```

```
[SwitchC] display ospf routing
```

```
OSPF Process 1 with Router ID 10.4.1.1
Routing Tables
```

```
Routing for Network
```

Destination	Cost	Type	NextHop	AdvRouter	Area
0.0.0.0/0	4	Inter	10.2.1.1	10.2.1.1	0.0.0.1
10.2.1.0/24	3	Transit	10.2.1.2	10.4.1.1	0.0.0.1
10.4.1.0/24	3	Stub	10.4.1.1	10.4.1.1	0.0.0.1

```
Total Nets: 3
```

```
Intra Area: 2 Inter Area: 1 ASE: 0 NSSA: 0
```

After this configuration, routing entries on the stub router are further reduced, containing only one default external route.

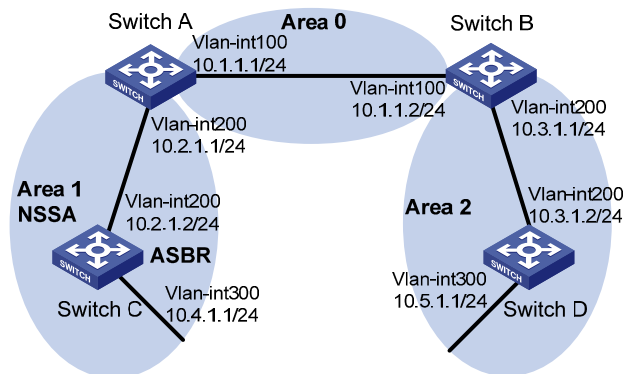
Configuring an OSPF NSSA area

Network requirements

Figure 42 shows an AS is split into three areas, where all switches run OSPF. Switch A and Switch B act as ABRs to forward routing information between areas.

Configure Area 1 as an NSSA area and configure Switch C as the ASBR to redistribute static routes into the AS.

Figure 42 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces.
2. Configure OSPF basic functions. (See "[Configuring OSPF basic functions](#)")
3. Configure Area 1 as an NSSA area:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] ospf
[SwitchA-ospf-1] area 1
[SwitchA-ospf-1-area-0.0.0.1] nssa default-route-advertise no-summary
[SwitchA-ospf-1-area-0.0.0.1] quit
[SwitchA-ospf-1] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] ospf
[SwitchC-ospf-1] area 1
[SwitchC-ospf-1-area-0.0.0.1] nssa
[SwitchC-ospf-1-area-0.0.0.1] quit
[SwitchC-ospf-1] quit
```

NOTE:

- If Switch C in the NSSA area wants to obtain routes to other areas within the AS, you must configure the **nssa** command with the keyword **default-route-advertise** on Switch A (an ABR) so Switch C can obtain a default route.
 - It is recommended to configure the **nssa** command with the keyword **no-summary** on Switch A to reduce the routing table size on NSSA switches. On other NSSA switches, you only need to configure the **nssa** command.
-

Display OSPF routing information on Switch C.

```
[SwitchC] display ospf routing
```

```
OSPF Process 1 with Router ID 10.4.1.1
Routing Tables
```

```

Routing for Network
Destination      Cost      Type      NextHop      AdvRouter      Area
0.0.0.0/0        65536     Inter     10.2.1.1     10.2.1.1       0.0.0.1
10.2.1.0/24      65535     Transit  10.2.1.2     10.4.1.1       0.0.0.1
10.4.1.0/24      3         Stub     10.4.1.1     10.4.1.1       0.0.0.1

```

```
Total Nets: 3
```

```
Intra Area: 2 Inter Area: 1 ASE: 0 NSSA: 0
```

4. Configure Switch C to redistribute static routes:

```
[SwitchC] ip route-static 3.1.3.1 24 10.4.1.2
```

```
[SwitchC] ospf
```

```
[SwitchC-ospf-1] import-route static
```

```
[SwitchC-ospf-1] quit
```

Display OSPF routing information on Switch D.

```
<SwitchD> display ospf routing
```

```
OSPF Process 1 with Router ID 10.5.1.1
```

```
Routing Tables
```

```

Routing for Network
Destination      Cost      Type      NextHop      AdvRouter      Area
10.2.1.0/24      22        Inter     10.3.1.1     10.3.1.1       0.0.0.2
10.3.1.0/24      10        Transit  10.3.1.2     10.3.1.1       0.0.0.2
10.4.1.0/24      25        Inter     10.3.1.1     10.3.1.1       0.0.0.2
10.5.1.0/24      10        Stub     10.5.1.1     10.5.1.1       0.0.0.2
10.1.1.0/24      12        Inter     10.3.1.1     10.3.1.1       0.0.0.2

```

```
Routing for ASEs
```

```

Destination      Cost      Type      Tag      NextHop      AdvRouter
3.1.3.0/24       1         Type2     1         10.3.1.1     10.2.1.1

```

```
Total Nets: 6
```

```
Intra Area: 2 Inter Area: 3 ASE: 1 NSSA: 0
```

The output shows that on Switch D an external route imported from the NSSA area.

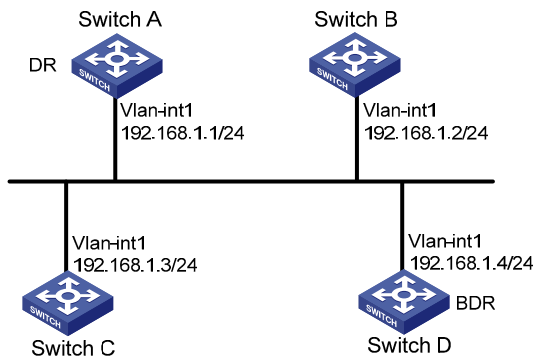
Configuring OSPF DR election

Network requirements

In [Figure 43](#), OSPF Switches A, B, C, and D reside on the same network segment.

Configure Switch A as the DR, and configure Switch C as the BDR.

Figure 43 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure OSPF basic functions:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] router id 1.1.1.1
[SwitchA] ospf
[SwitchA-ospf-1] area 0
[SwitchA-ospf-1-area-0.0.0.0] network 192.168.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] quit
[SwitchA-ospf-1] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] router id 2.2.2.2
[SwitchB] ospf
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 192.168.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] router id 3.3.3.3
[SwitchC] ospf
[SwitchC-ospf-1] area 0
[SwitchC-ospf-1-area-0.0.0.0] network 192.168.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] quit
[SwitchC-ospf-1] quit
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] router id 4.4.4.4
[SwitchD] ospf
[SwitchD-ospf-1] area 0
[SwitchD-ospf-1-area-0.0.0.0] network 192.168.1.0 0.0.0.255
[SwitchD-ospf-1-area-0.0.0.0] quit
```



```
[SwitchD-ospf-1] return
```

```
# Display OSPF neighbor information on Switch A.
```

```
[SwitchA] display ospf peer verbose
```

```
OSPF Process 1 with Router ID 1.1.1.1  
Neighbors
```

```
Area 0.0.0.0 interface 192.168.1.1(Vlan-interface1)'s neighbors
```

```
Router ID: 2.2.2.2          Address: 192.168.1.2          GR State: Normal
```

```
State: 2-Way Mode: None Priority: 1
```

```
DR: 192.168.1.4 BDR: 192.168.1.3 MTU: 0
```

```
Dead timer due in 38 sec
```

```
Neighbor is up for 00:01:31
```

```
Authentication Sequence: [ 0 ]
```

```
Router ID: 3.3.3.3          Address: 192.168.1.3          GR State: Normal
```

```
State: Full Mode: Nbr is Master Priority: 1
```

```
DR: 192.168.1.4 BDR: 192.168.1.3 MTU: 0
```

```
Dead timer due in 31 sec
```

```
Neighbor is up for 00:01:28
```

```
Authentication Sequence: [ 0 ]
```

```
Router ID: 4.4.4.4          Address: 192.168.1.4          GR State: Normal
```

```
State: Full Mode: Nbr is Master Priority: 1
```

```
DR: 192.168.1.4 BDR: 192.168.1.3 MTU: 0
```

```
Dead timer due in 31 sec
```

```
Neighbor is up for 00:01:28
```

```
Authentication Sequence: [ 0 ]
```

Switch D becomes the DR, and Switch C is the BDR.

3. Configure router priorities on interfaces:

```
# Configure Switch A.
```

```
[SwitchA] interface vlan-interface 1
```

```
[SwitchA-Vlan-interface1] ospf dr-priority 100
```

```
[SwitchA-Vlan-interface1] quit
```

```
# Configure Switch B.
```

```
[SwitchB] interface vlan-interface 1
```

```
[SwitchB-Vlan-interface1] ospf dr-priority 0
```

```
[SwitchB-Vlan-interface1] quit
```

```
# Configure Switch C.
```

```
[SwitchC] interface vlan-interface 1
```

```
[SwitchC-Vlan-interface1] ospf dr-priority 2
```

```
[SwitchC-Vlan-interface1] quit
```

```
# Display neighbor information on Switch D.
```

```
<SwitchD> display ospf peer verbose
```

```
OSPF Process 1 with Router ID 4.4.4.4  
Neighbors
```

```
Area 0.0.0.0 interface 192.168.1.4(Vlan-interface1)'s neighbors
Router ID: 1.1.1.1      Address: 192.168.1.1      GR State: Normal
State: Full Mode:Nbr is Slave Priority: 100
DR: 192.168.1.4 BDR: 192.168.1.3 MTU: 0
Dead timer due in 31 sec
Neighbor is up for 00:11:17
Authentication Sequence: [ 0 ]
```

```
Router ID: 2.2.2.2      Address: 192.168.1.2      GR State: Normal
State: Full Mode:Nbr is Slave Priority: 0
DR: 192.168.1.4 BDR: 192.168.1.3 MTU: 0
Dead timer due in 35 sec
Neighbor is up for 00:11:19
Authentication Sequence: [ 0 ]
```

```
Router ID: 3.3.3.3      Address: 192.168.1.3      GR State: Normal
State: Full Mode:Nbr is Slave Priority: 2
DR: 192.168.1.4 BDR: 192.168.1.3 MTU: 0
Dead timer due in 33 sec
Neighbor is up for 00:11:15
Authentication Sequence: [ 0 ]
```

The DR and BDR are not changed, because the new router priority settings do not take effect immediately.

4. Restart OSPF process:

Restart the OSPF process of Switch D.

```
<SwitchD> reset ospf 1 process
```

```
Warning : Reset OSPF process? [Y/N]:y
```

Display neighbor information on Switch D.

```
<SwitchD> display ospf peer verbose
```

```
OSPF Process 1 with Router ID 4.4.4.4
Neighbors
```

```
Area 0.0.0.0 interface 192.168.1.4(Vlan-interface1)'s neighbors
Router ID: 1.1.1.1      Address: 192.168.1.1      GR State: Normal
State: Full Mode: Nbr is Slave Priority: 100
DR: 192.168.1.1 BDR: 192.168.1.3 MTU: 0
Dead timer due in 39 sec
Neighbor is up for 00:01:40
Authentication Sequence: [ 0 ]
```

```
Router ID: 2.2.2.2      Address: 192.168.1.2      GR State: Normal
State: 2-Way Mode: None Priority: 0
DR: 192.168.1.1 BDR: 192.168.1.3 MTU: 0
Dead timer due in 35 sec
Neighbor is up for 00:01:44
Authentication Sequence: [ 0 ]
```

```

Router ID: 3.3.3.3          Address: 192.168.1.3      GR State: Normal
State: Full Mode: Nbr is Slave Priority: 2
DR: 192.168.1.1 BDR: 192.168.1.3 MTU: 0
Dead timer due in 39 sec
Neighbor is up for 00:01:41
Authentication Sequence: [ 0 ]

```

Switch A becomes the DR, and Switch C is the BDR.

The *full* neighbor state means an adjacency has been established. The *2-way* neighbor state means the two routers are not the DR or BDR, and they do not exchange LSAs.

Display OSPF interface information.

```
[SwitchA] display ospf interface
```

```

OSPF Process 1 with Router ID 1.1.1.1
Interfaces

```

```
Area: 0.0.0.0
```

IP Address	Type	State	Cost	Pri	DR	BDR
192.168.1.1	Broadcast	DR	1	100	192.168.1.1	192.168.1.3

```
[SwitchB] display ospf interface
```

```

OSPF Process 1 with Router ID 2.2.2.2
Interfaces

```

```
Area: 0.0.0.0
```

IP Address	Type	State	Cost	Pri	DR	BDR
		DROther				

The interface state *DROther* means the interface is not the DR/BDR.

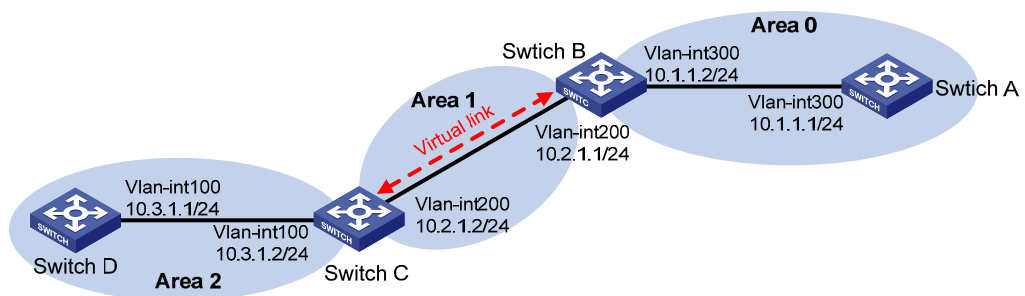
Configuring OSPF virtual links

Network requirements

In [Figure 44](#), Area 2 has no direct connection to Area 0, and Area 1 acts as the Transit Area to connect Area 2 to Area 0 via a configured virtual link between Switch B and Switch C.

After configuration, Switch B can learn routes to Area 2.

Figure 44 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure OSPF basic functions:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] ospf 1 router-id 1.1.1.1
[SwitchA-ospf-1] area 0
[SwitchA-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] ospf 1 router-id 2.2.2.2
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 10.1.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] area 1
[SwitchB-ospf-1-area-0.0.0.1] network 10.2.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.1] quit
[SwitchB-ospf-1] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] ospf 1 router-id 3.3.3.3
[SwitchC-ospf-1] area 1
[SwitchC-ospf-1-area-0.0.0.1] network 10.2.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.1] quit
[SwitchC-ospf-1] area 2
[SwitchC-ospf-1-area-0.0.0.2] network 10.3.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.2] quit
[SwitchC-ospf-1] quit
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] ospf 1 router-id 4.4.4.4
[SwitchD-ospf-1] area 2
[SwitchD-ospf-1-area-0.0.0.2] network 10.3.1.0 0.0.0.255
[SwitchD-ospf-1-area-0.0.0.2] quit
```

Display the OSPF routing table of Switch B.

```
[SwitchB] display ospf routing
```

```
OSPF Process 1 with Router ID 2.2.2.2
```

```
Routing Tables
```

```
Routing for Network
```

Destination	Cost	Type	NextHop	AdvRouter	Area
10.2.1.0/24	2	Transit	10.2.1.1	3.3.3.3	0.0.0.1
10.1.1.0/24	2	Transit	10.1.1.2	2.2.2.2	0.0.0.0

```
Total Nets: 2
```

```
Intra Area: 2 Inter Area: 0 ASE: 0 NSSA: 0
```

Area 0 has no direct connection to Area 2, so the routing table of Switch B has no route to Area 2.

3. Configure a virtual link:

Configure Switch B.

```
[SwitchB] ospf
[SwitchB-ospf-1] area 1
[SwitchB-ospf-1-area-0.0.0.1] vlink-peer 3.3.3.3
[SwitchB-ospf-1-area-0.0.0.1] quit
[SwitchB-ospf-1] quit
```

Configure Switch C.

```
[SwitchC] ospf 1
[SwitchC-ospf-1] area 1
[SwitchC-ospf-1-area-0.0.0.1] vlink-peer 2.2.2.2
[SwitchC-ospf-1-area-0.0.0.1] quit
```

Display the OSPF routing table of Switch B.

```
[SwitchB] display ospf routing
      OSPF Process 1 with Router ID 2.2.2.2
      Routing Tables

      Routing for Network
      Destination      Cost      Type      NextHop      AdvRouter      Area
      10.2.1.0/24      2         Transit  10.2.1.1     3.3.3.3        0.0.0.1
      10.3.1.0/24      5         Inter    10.2.1.2     3.3.3.3        0.0.0.0
      10.1.1.0/24      2         Transit  10.1.1.2     2.2.2.2        0.0.0.0
```

```
Total Nets: 3
```

```
Intra Area: 2  Inter Area: 1  ASE: 0  NSSA: 0
```

Switch B has learned the route 10.3.1.0/24 to Area 2.

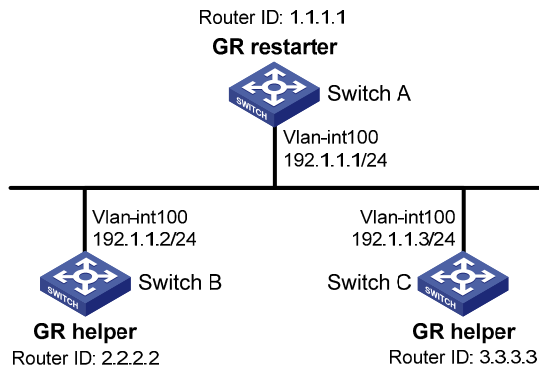
Configuring OSPF Graceful Restart

Network requirements

As shown in [Figure 45](#), Switch A, Switch B, and Switch C that belong to the same autonomous system and the same OSPF routing domain are GR capable.

Switch A acts as the non-IETF standard GR Restarter, whereas Switch B and Switch C are the GR Helpers and re-synchronize their LSDB with Switch A through OOB communication of GR.

Figure 45 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure OSPF basic functions:

Configure Switch A

```
<SwitchA> system-view
[SwitchA] router id 1.1.1.1
[SwitchA] ospf 100
[SwitchA-ospf-100] area 0
[SwitchA-ospf-100-area-0.0.0.0] network 192.1.1.0 0.0.0.255
[SwitchA-ospf-100-area-0.0.0.0] quit
```

Configure Switch B

```
<SwitchB> system-view
[SwitchB] router id 2.2.2.2
[SwitchB] ospf 100
[SwitchB-ospf-100] area 0
[SwitchB-ospf-100-area-0.0.0.0] network 192.1.1.0 0.0.0.255
[SwitchB-ospf-100-area-0.0.0.0] quit
```

Configure Switch C

```
<SwitchC> system-view
[SwitchC] router id 3.3.3.3
[SwitchC] ospf 100
[SwitchC-ospf-100] area 0
[SwitchC-ospf-100-area-0.0.0.0] network 192.1.1.0 0.0.0.255
[SwitchC-ospf-100-area-0.0.0.0] quit
```

3. Configure OSPF GR:

Configure Switch A as the non-IETF standard OSPF GR Restarter: enable the link-local signaling capability, the out-of-band re-synchronization capability, and non IETF standard Graceful Restart capability for OSPF process 100.

```
[SwitchA-ospf-100] enable link-local-signaling
[SwitchA-ospf-100] enable out-of-band-resynchronization
[SwitchA-ospf-100] graceful-restart
[SwitchA-ospf-100] return
```

Configure Switch B as the GR Helper: enable the link-local signaling capability and the out-of-band re-synchronization capability for OSPF process 100.

```
[SwitchB-ospf-100] enable link-local-signaling
[SwitchB-ospf-100] enable out-of-band-resynchronization
# Configure Switch C as the GR Helper: enable the link-local signaling capability and the
out-of-band re-synchronization capability for OSPF process 100.
[SwitchC-ospf-100] enable link-local-signaling
[SwitchC-ospf-100] enable out-of-band-resynchronization
```

4. Verify the configuration:

After the configurations on Switch A, Switch B, and Switch C are completed and the switches are running steadily, enable OSPF Graceful Restart event debugging and then restart the OSPF process using GR on Switch A.

```
<SwitchA> debugging ospf event graceful-restart
<SwitchA> terminal monitor
<SwitchA> terminal debugging
<SwitchA> reset ospf 100 process graceful-restart
Warning : Reset OSPF process? [Y/N]:y
%Dec 12 09:36:12:500 2006 SwitchA RM/3/RMLOG:OSPF-NBRCHANGE: Process 100, Neighbour
192.1.1.1(Vlan100) from Full to Down
OSPF 100: Intf 192.1.1.1 Rcv InterfaceDown State BackupDR -> Down.
OSPF 100 nonstandard GR Started for OSPF Router
OSPF 100 notify RM that OSPF process will enter GR.
OSPF 100 created GR wait timer, timeout interval is 40(s).
OSPF 100 created GR Interval timer, timeout interval is 120(s).
OSPF 100: Intf 192.1.1.1 Rcv InterfaceUp State Down -> Waiting.
OSPF 100: Intf 192.1.1.1 Rcv BackupSeen State Waiting -> BackupDR.
OSPF 100 created OOB Progress timer for neighbor 192.1.1.2.
OSPF 100 restarted OOB Progress timer for neighbor 192.1.1.2.
OSPF 100 restarted OOB Progress timer for neighbor 192.1.1.2.
%Oct 22 09:36:12:566 2008 SwitchA RM/3/RMLOG:OSPF-NBRCHANGE: Process 100, Neighbour
192.1.1.2(Vlan100) from Loading to Full
OSPF 100 restarted OOB Progress timer for neighbor 192.1.1.2.
OSPF 100 deleted OOB Progress timer for neighbor 192.1.1.2.
OSPF 100 Gr Wait Timeout timer fired.
OSPF 100 deleted GR wait timer.
OSPF 100 deleted GR Interval timer.
OSPF 100 GR Completed for OSPF Router
OSPF 100 notified RM that OSPF process left GR.
RM notified that all protocol left GR.
OSPF 100 started flushing STALE LSA after all protocol left GR.
OSPF 100: Flush Stale Area LSAs
OSPF 100: Start Flush Stale ASE + NSSA LSAs
OSPF 100: End Flush Stale ASE + NSSA LSAs
```

Switch A completes GR with the help of Switch B.

Configuring route filtering

Network requirements

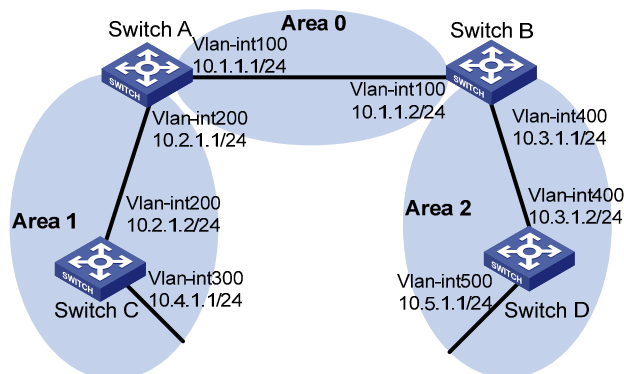
As shown in [Figure 46](#):

- All the switches in the network run OSPF. The AS is divided into three areas.
- Switch A and Switch B work as ABRs.

Configure Switch C as an ASBR to redistribute external routes (static routes), and configure a filter policy on Switch C to filter out redistributed route 3.1.3.0/24.

Configure a routing policy on Switch A to filter route 10.5.1.0/24.

Figure 46 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure OSPF basic functions. (See "[Configuring OSPF basic functions](#)")
3. Configure OSPF to redistribute routes:

On Switch C, configure a static route destined for network 3.1.1.0/24.

```
<SwitchC> system-view
[SwitchC] ip route-static 3.1.1.0 24 10.4.1.2
```

On Switch C, configure a static route destined for network 3.1.2.0/24.

```
[SwitchC] ip route-static 3.1.2.0 24 10.4.1.2
```

On Switch C, configure a static route destined for network 3.1.3.0/24.

```
[SwitchC] ip route-static 3.1.3.0 24 10.4.1.2
```

On Switch C, configure OSPF to redistribute static routes.

```
[SwitchC] ospf 1
[SwitchC-ospf-1] import-route static
[SwitchC-ospf-1] quit
```

Display the OSPF routing table of Switch A.

```
<SwitchA> display ip routing-table
```

```
Routing Tables: Public
```

```
Destinations : 12      Routes : 12
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
------------------	-------	-----	------	---------	-----------

3.1.1.0/24	O_ASE	150	1	10.2.1.2	Vlan200
3.1.2.0/24	O_ASE	150	1	10.2.1.2	Vlan200
3.1.3.0/24	O_ASE	150	1	10.2.1.2	Vlan200
10.1.1.0/24	Direct	0	0	10.1.1.1	Vlan200
10.1.1.1/32	Direct	0	0	127.0.0.1	InLoop0
10.2.1.0/24	Direct	0	0	10.2.1.1	Vlan200
10.2.1.1/32	Direct	0	0	127.0.0.1	InLoop0
10.3.1.0/24	OSPF	10	4	10.1.1.2	Vlan100
10.4.1.0/24	OSPF	10	13	10.2.1.2	Vlan200
10.5.1.0/24	OSPF	10	14	10.1.1.2	Vlan100
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0

4. On Switch C, filter out route 3.1.3.0/24:

Configure the IPv4 prefix list.

```
[SwitchC] ip ip-prefix prefix1 index 1 deny 3.1.3.0 24
[SwitchC] ip ip-prefix prefix1 index 2 permit 3.1.1.0 24
[SwitchC] ip ip-prefix prefix1 index 3 permit 3.1.2.0 24
```

Reference the prefix list to filter out route 3.1.3.0/24.

```
[SwitchC] ospf 1
[SwitchC-ospf-1] filter-policy ip-prefix prefix1 export static
```

Display the OSPF routing table of Switch A.

```
<SwitchA> display ip routing-table
```

Routing Tables: Public

Destinations : 11 Routes : 11

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
3.1.1.0/24	O_ASE	150	1	10.2.1.2	Vlan200
3.1.2.0/24	O_ASE	150	1	10.2.1.2	Vlan200
10.1.1.0/24	Direct	0	0	10.1.1.1	Vlan100
10.1.1.1/32	Direct	0	0	127.0.0.1	InLoop0
10.2.1.0/24	Direct	0	0	10.2.1.1	Vlan200
10.2.1.1/32	Direct	0	0	127.0.0.1	InLoop0
10.3.1.0/24	OSPF	10	4	10.1.1.2	Vlan100
10.4.1.0/24	OSPF	10	13	10.2.1.2	Vlan200
10.5.1.0/24	OSPF	10	14	10.1.1.2	Vlan100
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0

The route destined for network 3.1.3.0/24 is filtered out.

5. On Switch A, filter out the route 10.5.1.1/24:

Configure the ACL on Switch A.

```
<SwitchA> system-view
[SwitchA] acl number 2000
[SwitchA-acl-basic-2000] rule 0 deny source 10.5.1.0 0.0.0.255
[SwitchA-acl-basic-2000] rule 1 permit source any
[SwitchA-acl-basic-2000] quit
```

Use the ACL to filter route 10.5.1.0/24.

```
[SwitchA] ospf 1
[SwitchA-ospf-1] filter-policy 2000 import
[SwitchA-ospf-1] quit
# Display the OSPF routing table of Switch A.
[SwitchA] display ip routing-table
Routing Tables: Public
                Destinations : 10          Routes : 10
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
3.1.1.0/24	O_ASE	150	1	10.2.1.2	Vlan200
3.1.2.0/24	O_ASE	150	1	10.2.1.2	Vlan200
10.1.1.0/24	Direct	0	0	10.1.1.1	Vlan100
10.1.1.1/32	Direct	0	0	127.0.0.1	InLoop0
10.2.1.0/24	Direct	0	0	10.2.1.1	Vlan200
10.2.1.1/32	Direct	0	0	127.0.0.1	InLoop0
10.3.1.0/24	OSPF	10	4	10.1.1.2	Vlan100
10.4.1.0/24	OSPF	10	13	10.2.1.2	Vlan200
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0

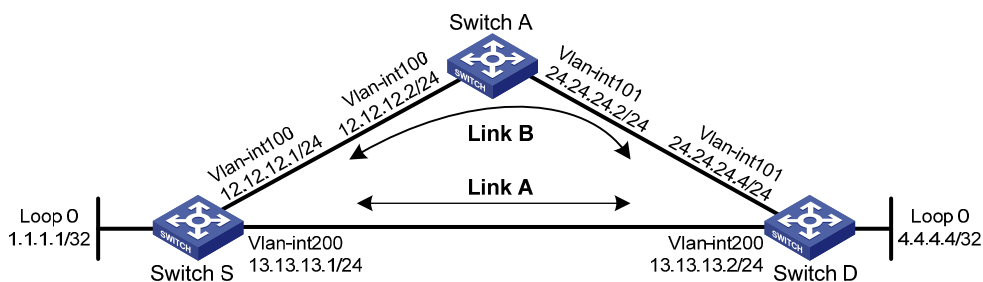
The route destined for 10.5.1.1/24 is filtered out.

Configuring OSPF FRR

Network requirements

Switch S, Switch A, and Switch D belong to the same OSPF domain as illustrated in [Figure 47](#). Configure OSPF FRR so that when the link between Switch S and Switch D fails, traffic can be switched to Link B immediately.

Figure 47 Network diagram



Configuration procedure

1. Configure IP addresses for the interfaces on each switch and configure OSPF.
Follow [Figure 47](#) to configure the IP address and subnet mask of each interface on the switches. (Details not shown.)
Configure OSPF on the switches, ensuring that Switch S, Switch A, and Switch D can communicate with each other at Layer 3. (Details not shown.)
2. Configure OSPF FRR:

You can enable OSPF FRR to either automatically calculate a backup next hop, or to designate a backup next hop by using a routing policy.

Method I: Enable OSPF FRR to automatically calculate the backup next hop.

Configure Switch S.

```
<SwitchS> system-view
[SwitchS] bfd echo-source-ip 1.1.1.1
[SwitchS] ospf 1
[SwitchS-ospf-1] fast-reroute auto
[SwitchS-ospf-1] quit
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] bfd echo-source-ip 4.4.4.4
[SwitchD] ospf 1
[SwitchD-ospf-1] fast-reroute auto
[SwitchD-ospf-1] quit
```

Method II: Enable OSPF FRR to designate a backup next hop by using a routing policy.

Configure Switch S.

```
<SwitchS> system-view
[SwitchS] bfd echo-source-ip 1.1.1.1
[SwitchS] ip ip-prefix abc index 10 permit 4.4.4.4 32
[SwitchS] route-policy frr permit node 10
[SwitchS-route-policy] if-match ip-prefix abc
[SwitchS-route-policy] apply fast-reroute backup-interface vlan-interface 100
backup-nexthop 12.12.12.2
[SwitchS-route-policy] quit
[SwitchS] ospf 1
[SwitchS-ospf-1] fast-reroute route-policy frr
[SwitchS-ospf-1] quit
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] bfd echo-source-ip 4.4.4.4
[SwitchD] ip ip-prefix abc index 10 permit 1.1.1.1 32
[SwitchD] route-policy frr permit node 10
[SwitchD-route-policy] if-match ip-prefix abc
[SwitchD-route-policy] apply fast-reroute backup-interface vlan-interface 101
backup-nexthop 24.24.24.2
[SwitchD-route-policy] quit
[SwitchD] ospf 1
[SwitchD-ospf-1] fast-reroute route-policy frr
[SwitchD-ospf-1] quit
```

3. Verify the configuration:

Display route 4.4.4.4/32 on Switch S and you can view the backup next hop information.

```
[SwitchS] display ip routing-table 4.4.4.4 verbose
Routing Table : Public
Summary Count : 1
```

```
Destination: 4.4.4.4/32
```

```

    Protocol: OSPF                Process ID: 1
    Preference: 10                Cost: 1
    IpPrecedence:                QoSLeId:
    NextHop: 13.13.13.2          Interface: Vlan-interface200
    BkNextHop: 12.12.12.2       BkInterface: Vlan-interface100
    RelyNextHop: 0.0.0.0        Neighbor : 0.0.0.0
    Tunnel ID: 0x0              Label: NULL
    BKTunnel ID: 0x0           BKLabel: NULL
    State: Active Adv          Age: 00h01m27s
    Tag: 0

```

Display route 1.1.1.1/32 on Switch D. You can find the backup next hop information.

```

[SwitchD] display ip routing-table 1.1.1.1 verbose
Routing Table : Public
Summary Count : 1

```

```

    Destination: 1.1.1.1/32
    Protocol: OSPF                Process ID: 1
    Preference: 10                Cost: 1
    IpPrecedence:                QoSLeId:
    NextHop: 13.13.13.1          Interface: Vlan-interface200
    BkNextHop: 24.24.24.2       BkInterface: Vlan-interface101
    RelyNextHop: 0.0.0.0        Neighbor : 0.0.0.0
    Tunnel ID: 0x0              Label: NULL
    BKTunnel ID: 0x0           BKLabel: NULL
    State: Active Adv          Age: 00h01m27s
    Tag: 0

```

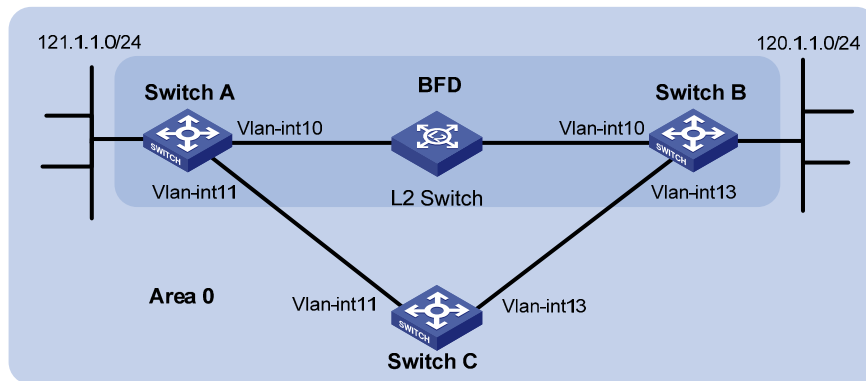
Configuring BFD for OSPF

Network requirements

As shown in [Figure 48](#), OSPF is enabled on Switch A, Switch B and Switch C that are reachable to each other at the network layer.

After the link over which Switch A and Switch B communicate through a Layer 2 switch fails, BFD can quickly detect the failure and notify OSPF of the failure. Switch A and Switch B then communicate through Switch C.

Figure 48 Network diagram



Device	Interface	IP address	Device	Interface	IP address
Switch A	Vlan-int10	10.1.0.102/24	Switch B	Vlan-int10	10.1.0.100/24
	Vlan-int11	11.1.1.1/24		Vlan-int13	13.1.1.1/24
Switch C	Vlan-int11	11.1.1.2/24			
	Vlan-int13	13.1.1.2/24			

Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure OSPF basic functions:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] ospf
[SwitchA-ospf-1] area 0
[SwitchA-ospf-1-area-0.0.0.0] network 10.1.0.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] network 11.1.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] network 121.1.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] quit
[SwitchA-ospf-1] quit
[SwitchA] interface vlan 11
[SwitchA-Vlan-interface11] ospf cost 2
[SwitchA-Vlan-interface11] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] ospf
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 10.1.0.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] network 13.1.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] network 120.1.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] quit
[SwitchB] interface vlan-interface 13
[SwitchB-Vlan-interface13] ospf cost 2
[SwitchB-Vlan-interface13] quit
```

Configure Switch C.

```
<SwitchC> system-view
```

```
[SwitchC] ospf
[SwitchC-ospf-1] area 0
[SwitchC-ospf-1-area-0.0.0.0] network 11.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] network 13.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] quit
[SwitchC-ospf-1] quit
```

3. Configure BFD:

Enable BFD on Switch A and configure BFD parameters.

```
[SwitchA] bfd session init-mode active
[SwitchA] interface vlan-interface 10
[SwitchA-Vlan-interface10] ospf bfd enable
[SwitchA-Vlan-interface10] bfd min-transmit-interval 500
[SwitchA-Vlan-interface10] bfd min-receive-interval 500
[SwitchA-Vlan-interface10] bfd detect-multiplier 7
[SwitchA-Vlan-interface10] quit
[SwitchA] quit
```

Enable BFD on Switch B and configure BFD parameters.

```
[SwitchB] bfd session init-mode active
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] ospf bfd enable
[SwitchB-Vlan-interface10] bfd min-transmit-interval 500
[SwitchB-Vlan-interface10] bfd min-receive-interval 500
[SwitchB-Vlan-interface10] bfd detect-multiplier 6
```

4. Verify the configuration:

The following operations are performed on Switch A. The operations on Switch B and Switch C are similar. (Details not shown.)

Display the BFD information of Switch A.

```
<SwitchA> display bfd session
Total Session Num: 1          Init Mode: Active
Session Working Under Ctrl Mode:
LD/RD      SourceAddr      DestAddr      State Holdtime Interface
3/1        10.1.0.102      10.1.0.100   Up    1700ms  vlan10
```

Display routes to 120.1.1.0/24 on Switch A, and you can see that Switch A communicates with Switch B through the Layer 2 switch.

```
<SwitchA> display ip routing-table 120.1.1.0 verbose
Routing Table : Public
Summary Count : 2
  Destination: 120.1.1.0/24
    Protocol: OSPF          Process ID: 0
    Preference: 0          Cost: 2
    IpPrecedence:          QoSLeId:
      NextHop: 192.168.0.100  Interface: Vlan-interface10
      BkNextHop: 0.0.0.0      BkInterface:
      RelyNextHop: 0.0.0.0    Neighbor : 0.0.0.0
      Tunnel ID: 0x0          Label: NULL
      BKTunnel ID: 0x0       BKLabel: NULL
      State: Active Adv      Age: 00h58m10s
```

```

Tag: 0
Destination: 120.1.1.0/24
  Protocol: OSPF          Process ID: 1
  Preference: 10         Cost: 4
IpPrecedence:           QoSLeId:
  NextHop: 10.1.1.100    Interface: Vlan-interface11
  BkNextHop: 0.0.0.0     BkInterface:
  RelyNextHop: 0.0.0.0   Neighbor : 0.0.0.0
  Tunnel ID: 0x0         Label: NULL
  BKTunnel ID: 0x0      BKLabel: NULL
  State: Invalid Adv     Age: 00h58m05s
Tag: 0

```

Enable BFD debugging on Switch A.

```

<SwitchA> debugging bfd scm
<SwitchA> debugging bfd event
<SwitchA> debugging ospf event
<SwitchA> terminal debugging

```

After the link over which Switch A and Switch B communicates through the Layer 2 switch fails, Switch A can quickly detect the changes on Switch B.

```

%Nov 12 18:34:48:823 2005 SwitchA BFD/5/LOG: Sess[10.1.0.102/10.1.0.100, vlan10],
Sta : UP->DOWN, Diag: 1
%Nov 12 18:34:48:824 2005 SwitchA RM/4/RMLOG:OSPF-NBRCHANGE: Process 1, Neighbour
10.1.0.102 (vlan10) from Full to Down
*0.50673825 SwitchA BFD/8/SCM:Sess[10.1.0.102/10.1.0.100, vlan10], Oper: Reset
*0.50673825 SwitchA BFD/8/EVENT:Send sess-down Msg, [Src:10.1.0.102, Dst:10.1.0.100,
vlan10] Protocol: OSPF

*0.50673826 SwitchA RM/7/RMDEBUG:OSPF-BFD: Message Type rcv BFD down, Connect Type
direct-connect, Src IP Address 10.1.0.102, Src IFIndex 5, Dst IP Address 10.1.0.100

*0.50673827 SwitchA RM/7/RMDEBUG:OSPF-BFD: Message Type delete session, Connect Type
direct-connect, Src IP Address 10.1.0.102, Src IFIndex 5, Dst IP Address 10.1.0.100
OSPF 1: Nbr 10.1.0.100 Rcv KillNbr State Full -> Down.
*0.50673829 SwitchA BFD/8/EVENT:Receive Delete-sess, [Src:10.1.0.102,
Dst:10.1.0.100, vlan10], Direct, Proto:OSPF
*0.50673830 SwitchA BFD/8/SCM:Sess[10.1.0.102/10.1.0.100, vlan10], Oper: Del
application(OSPF)
*0.50673831 SwitchA BFD/8/SCM:No application in session, delete
session[10.1.0.102/10.1.0.100, vlan10]
*0.50673831 SwitchA BFD/8/SCM:Sess[10.1.0.102/10.1.0.100, vlan10], Oper: Delete
*0.50673832 SwitchA BFD/8/SCM>Delete send-packet timer
*0.50673833 SwitchA BFD/8/SCM>Delete session entry
*0.50673833 SwitchA BFD/8/SCM>Delete session from IP hash table
*0.50673834 SwitchA BFD/8/SCM>Delete session from bfd interface
*0.50673834 SwitchA BFD/8/SCM:No session under bfd-int[vlan10] with default
configuration, delete bfd-if
*0.50673835 SwitchA BFD/8/SCM:Bfd-if[vlan10], Oper: Delete
*0.50673840 SwitchA BFD/8/SCM:No bfd session exists, stop receiving any bfd packets

```

Display the BFD information of Switch A.

The BFD session between Switch A and Switch B is deleted and no information is output.

```
<SwitchA> display bfd session
```

Display routes to 120.1.1.0/24 on Switch A, and you can see that Switch A communicates with Switch B through Switch C.

```
<SwitchA> display ip routing-table 120.1.1.0 verbose
```

```
Routing Table : Public
```

```
Summary Count : 2
```

```
Destination: 120.1.1.0/24
```

```
Protocol: OSPF Process ID: 1
```

```
Preference: 10 Cost: 4
```

```
IpPrecedence: QoSLeId:
```

```
NextHop: 10.1.1.100 Interface: Vlan-interface11
```

```
BkNextHop: 0.0.0.0 BkInterface:
```

```
RelyNextHop: 0.0.0.0 Neighbor : 0.0.0.0
```

```
Tunnel ID: 0x0 Label: NULL
```

```
BKTunnel ID: 0x0 BKLabel: NULL
```

```
State: Active Adv Age: 00h58m10s
```

```
Tag: 0
```

```
Destination: 120.1.1.0/24
```

```
Protocol: OSPF Process ID: 0
```

```
Preference: 0 Cost: 2
```

```
IpPrecedence: QoSLeId:
```

```
NextHop: 192.168.0.100 Interface: Vlan-interface10
```

```
BkNextHop: 0.0.0.0 BkInterface:
```

```
RelyNextHop: 0.0.0.0 Neighbor : 0.0.0.0
```

```
Tunnel ID: 0x0 Label: NULL
```

```
BKTunnel ID: 0x0 BKLabel: NULL
```

```
State: Invalid Adv Age: 00h58m05s
```

```
Tag: 0
```

Troubleshooting OSPF configuration

No OSPF neighbor relationship established

Symptom

No OSPF neighbor relationship can be established.

Analysis

If the physical link and lower layer protocols work well, check OSPF parameters configured on interfaces. Two neighbors must have the same parameters, such as the area ID, network segment, and mask (a P2P or virtual link may have different network segments and masks).

Solution

1. Display OSPF neighbor information using the **display ospf peer** command.
2. Display OSPF interface information using the **display ospf interface** command.
3. Ping the neighbor router's IP address to check connectivity.
4. Check OSPF timers. The dead interval on an interface must be at least four times the hello interval.

5. On an NBMA network, using the **peer ip-address** command to specify the neighbor manually is required.
6. On an NBMA or a broadcast network, at least one connected interface must have a router priority higher than 0.

Incorrect routing information

Symptom

OSPF cannot find routes to other areas.

Analysis

The backbone area must maintain connectivity to all other areas. If a router connects to more than one area, at least one area must be connected to the backbone. The backbone cannot be configured as a Stub area.

In a Stub area, all routers cannot receive external routes, and all interfaces connected to the Stub area must belong to the Stub area.

Solution

1. Use the **display ospf peer** command to display neighbors.
2. Use the **display ospf interface** command to display OSPF interface information.
3. Use the **display ospf lsdb** command to display the LSDB to check its integrity.
4. Display information about area configuration using the **display current-configuration configuration ospf** command. If more than two areas are configured, at least one area is connected to the backbone.
5. In a Stub area, all routers attached are configured with the **stub** command. In an NSSA area, all routers attached are configured with the **nssa** command.
6. If a virtual link is configured, use the **display ospf vlink** command to check the state of the virtual link.

Configuring IS-IS

Hardware compatibility

The HP 5500 SI Switch Series does not support IS-IS.

IS-IS overview

Intermediate System-to-Intermediate System (IS-IS) is a dynamic routing protocol designed by the International Organization for Standardization (ISO) to operate on the connectionless network protocol (CLNP).

The IS-IS routing protocol was modified and extended in RFC 1195 by the International Engineer Task Force (IETF) for application in both TCP/IP and OSI reference models, and the new one is named "Integrated IS-IS" or "Dual IS-IS".

IS-IS is an Interior Gateway Protocol (IGP) used within an Autonomous System. It adopts the Shortest Path First (SPF) algorithm for route calculation.

The term "router" in this chapter refers to both routers and Layer 3 switches.

Basic concepts

IS-IS terminology

- **Intermediate system (IS)**—Similar to a router in TCP/IP, it is the basic unit in IS-IS to generate and propagate routing information. In the following text, an IS refers to a router.
- **End system (ES)**—Refers to a host system in TCP/IP. ISO defines the ES-IS protocol for communication between an ES and an IS. An ES does not participate in the IS-IS processing.
- **Routing domain (RD)**—A group of ISs exchanges routing information with each other using the same routing protocol in a routing domain.
- **Area**—A unit in a routing domain. The IS-IS protocol allows a routing domain to be divided into multiple areas.
- **Link State Database (LSDB)**—All link states in the network forms the LSDB. Each IS has at least one LSDB. The IS uses the SPF algorithm and LSDB to generate its own routes.
- **Link State Protocol Data Unit (LSPDU) or Link State Packet (LSP)**—Each IS can generate an LSP, which contains all the link state information of the IS.
- **Network Protocol Data Unit (NPDU)**—A network layer protocol packet in OSI, which is equivalent to an IP packet in TCP/IP.
- **Designated IS**—On a broadcast network, the designated router is also known as the "designated IS".
- **Network service access point (NSAP)**—An NSAP is an OSI network layer address. It identifies an abstract network service access point and describes the network address in the OSI reference model.

IS-IS address format

- NSAP

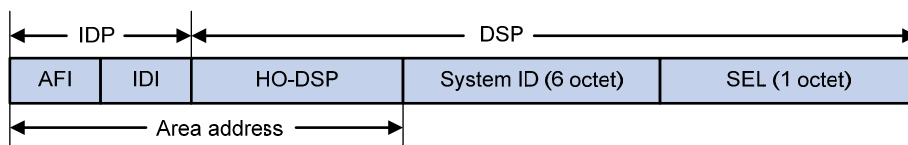
As shown in [Figure 49](#), an NSAP address consists of the Initial Domain Part (IDP) and the Domain Specific Part (DSP). The IDP is equal to the network ID of an IP address, and the DSP is equal to the subnet and host ID.

The IDP includes the Authority and Format Identifier (AFI) and the Initial Domain Identifier (IDI).

The DSP includes the High Order Part of DSP (HO-DSP), System ID, and SEL, where the HO-DSP identifies the area, the System ID identifies the host, and the SEL identifies the type of service.

The IDP and DSP are variable in length. The length of an NSAP address varies from 8 bytes to 20 bytes.

Figure 49 NSAP address format



- Area address

The area address comprises the IDP and the HO-DSP of the DSP, which identify the area and the routing domain. Different routing domains cannot have the same area address.

Typically, a router only needs one area address, and all nodes in the same routing domain must share the same area address. However, a router can have a maximum of three area addresses to support smooth area merging, partitioning, and switching.

- System ID

A system ID identifies a host or router uniquely. It has a fixed length of 48 bits (6 bytes).

The system ID of a device can be generated from the Router ID. For example, a router uses the IP address 168.10.1.1 of Loopback 0 as the Router ID. The system ID in IS-IS can be obtained in the following ways:

- Extend each decimal number of the IP address to 3 digits by adding 0s from the left, like 168.010.001.001;
- Divide the extended IP address into 3 sections with 4 digits in each section to get the system ID 1680.1000.1001.

If you use other methods for defining a system ID, always make sure that it can uniquely identify a host or router.

- SEL

The NSAP Selector (SEL), or the N-SEL, is similar to the protocol identifier in IP. Different transport layer protocols correspond to different SELs. All SELs in IP are 00.

- Routing method

Because the area information is identified in IS-IS addresses, a Level-1 router can easily identify packets destined to other areas.

- A Level-1 router makes routing decisions based on the system ID. If the destination is not in the area, the packet is forwarded to the nearest Level-1-2 router.
- A Level-2 router routes packets across areas according to the area address.

NET

A network entity title (NET) indicates the network layer information of an IS, and does not include transport layer information. It is a special NSAP address with the SEL being 0. The length of the NET is equal to the NSAP, and is in the range of 8 bytes to 20 bytes.

A NET comprises the following parts:

- **Area ID**—Its length is in the range of 1 to 13 bytes.
- **System ID**—A system ID uniquely identifies a host or router in the area and has a fixed 6-byte length.
- **SEL**—It has a value of 0 and a fixed 1-byte length.

For example, a NET is ab.cdef.1234.5678.9abc.00, where, area ID is ab.cdef, system ID is 1234.5678.9abc, and SEL is 00.

Typically, a router only needs one NET, but it can have a maximum of three NETs for smooth area merging and partitioning. When you configure multiple NETs, ensure their system IDs are the same.

IS-IS area

Two-level hierarchy

IS-IS has a two-level hierarchy to support large scale networks. A large scale routing domain is divided into multiple Areas. Typically, a Level-1 router is deployed within an area, a Level-2 router is deployed between areas, and a Level-1-2 router is deployed between Level-1 and Level-2 routers.

Level-1 and Level-2

- **Level-1 router**—A Level-1 router establishes neighbor relationships with Level-1 and Level-1-2 routers in the same area. The LSDB maintained by the Level-1 router contains the local area routing information. It directs the packets destined for an outside area to the nearest Level-1-2 router.
- **Level-2 router**—A Level-2 router establishes neighbor relationships with the Level-2 and Level-1-2 routers in the same or in different areas. It maintains a Level-2 LSDB containing inter-area routing information. All the Level-2 and Level-1-2 routers must be contiguous to form the backbone of a routing domain.
- **Level-1-2 router**—A router with both Level-1 and Level-2 router functions is a Level-1-2 router. It can establish Level-1 neighbor relationships with the Level-1 and Level-1-2 routers in the same area, or establish Level-2 neighbor relationships with the Level-2 and Level-1-2 routers in different areas. A Level-1 router must be connected to other areas through a Level-1-2 router. The Level-1-2 router maintains two LSDBs, where the Level-1 LSDB is for routing within the area, and the Level-2 LSDB is for routing between areas.

The Level-1 routers in different areas cannot establish neighbor relationships.

The neighbor relationship establishment of Level-2 routers has nothing to do with area.

Figure 50 shows an IS-IS network topology. Area 1 comprises a set of Level-2 routers and is the backbone. The other four areas are non-backbone areas connected to the backbone through Level-1-2 routers.

Figure 50 IS-IS topology 1

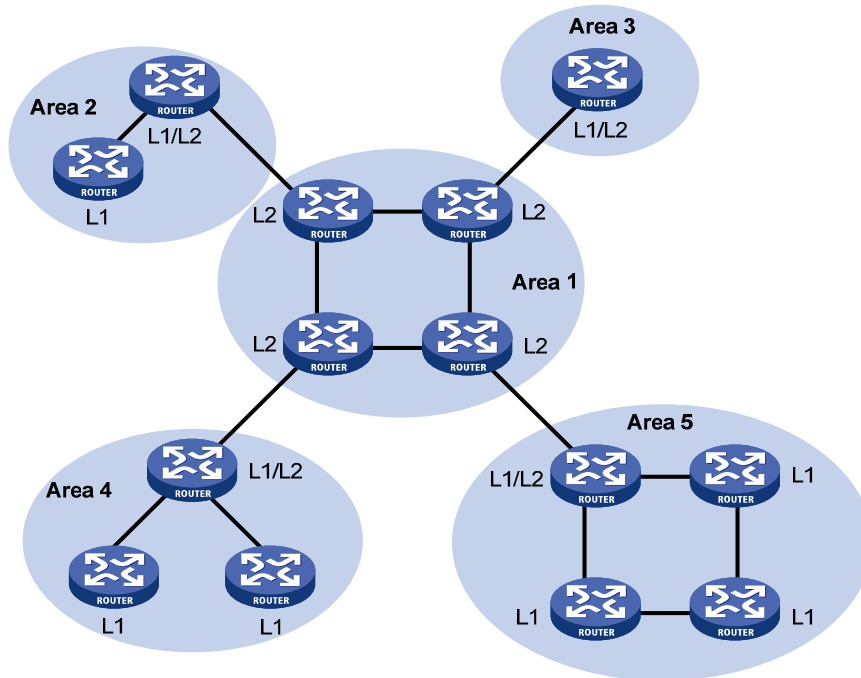
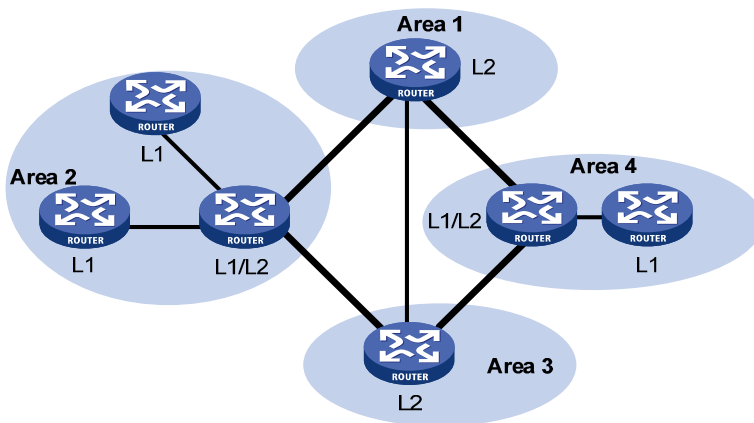


Figure 51 is another IS-IS topology. The Level-1-2 routers connect to the Level-1 and Level-2 routers, and form the IS-IS backbone together with the Level-2 routers. No area is defined as the backbone in this topology. The backbone comprises all contiguous Level-2 and Level-1-2 routers, which can reside in different areas.

Figure 51 IS-IS topology 2



NOTE:

The IS-IS backbone does not need to be a specific area.

Both the Level-1 and Level-2 routers use the SPF algorithm to generate the shortest path tree (SPT).

Route leaking

An IS-IS routing domain is comprised of only one Level-2 area and multiple Level-1 areas. A Level-1 area consists of a group of Level-1 routers, and is connected with a Level-2 area rather than other Level-1 areas.

The routing information of a Level-1 area is sent to the Level-2 area through the Level-1-2 router; therefore, the Level-2 router knows the routing information of the entire IS-IS routing domain. But the Level-1-2 router does not share the information of other Level-1 areas and the Level-2 area with the Level-1 area by default.

Because a Level-1 router simply sends packets destined for other areas to the nearest Level-1-2 router, the best paths may not be selected. To resolve this problem, route leaking was introduced. A Level-2 router can advertise Level-2 routing information to a specified Level-1 area. By having the routing information of other areas, a Level-1 router in the area can make a better routing decision for a packet to another area.

IS-IS network type

Network type

IS-IS supports the following network types:

- Broadcast network, such as Ethernet and Token-Ring
- Point-to-point network, such as PPP and HDLC

DIS and pseudonodes

On an IS-IS broadcast network, a router is elected as the Designated Intermediate System (DIS).

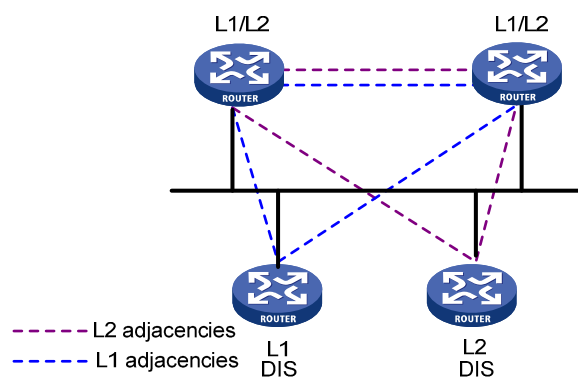
The Level-1 and Level-2 DISs are elected. You can assign different priorities to different level DIS elections. The higher a router's priority is, the more likely the router becomes the DIS. If multiple routers with the same highest DIS priority exist, the one with the highest SNPA (Subnetwork Point of Attachment) address (MAC address on a broadcast network) will be elected. A router can be the DIS for different levels.

IS-IS DIS election differs from OSPF DIS election in the following ways:

- A router with priority 0 can also participate in the DIS election.
- When a router is added to the network and becomes the new DIS, an LDP flooding process is triggered.

As shown in [Figure 52](#), the same level routers on a network, including non-DIS routers, establish adjacencies with each other.

Figure 52 DIS in the IS-IS broadcast network



The DIS creates and updates pseudonodes, as well as generates their LSPs, to describe all routers on the network.

A pseudonode represents a virtual node on the broadcast network. It is not a real router. In IS-IS, it is identified by the system ID of the DIS and a one-byte Circuit ID (a non zero value).

Using pseudonodes can reduce the resources consumed by SPF and simplify network topology.

NOTE:

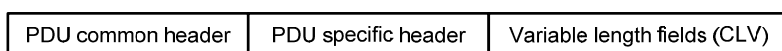
On IS-IS broadcast networks, all routers are adjacent with each other. However, the DIS is responsible for the synchronization of their LSDBs.

IS-IS PDU format

PDU header format

IS-IS packets are encapsulated into link layer frames. The Protocol Data Unit (PDU) consists of two parts, the headers and the variable length fields. The headers comprise the PDU common header and the PDU specific header. All PDUs have the same PDU common header. The specific headers vary by PDU type.

Figure 53 PDU format



Common header format

Figure 54 PDU common header format

				No. of Octets
Intradomain routing protocol discriminator				1
Length indicator				1
Version/Protocol ID extension				1
ID length				1
R	R	R	PDU type	1
Version				1
Reserved				1
Maximum area address				1

Major fields of the PDU common header are as follows:

- **Intradomain routing protocol discriminator**—Set to 0x83.
- **Length indicator**—Length of the PDU header in bytes, including both common and specific headers.
- **Version/Protocol ID extension**—Set to 1(0x01).
- **ID length**—Length of the NSAP address and NET ID.
- **R (Reserved)**—Set to 0.
- **PDU type**—See [Table 4](#).
- **Version**—Set to 1(0x01).
- **Maximum area address**—Maximum number of area addresses supported.

Table 4 PDU type

Type	PDU Type	Acronym
15	Level-1 LAN IS-IS hello PDU	L1 LAN IIH
16	Level-2 LAN IS-IS hello PDU	L2 LAN IIH
17	Point-to-Point IS-IS hello PDU	P2P IIH

Type	PDU Type	Acronym
18	Level-1 Link State PDU	L1 LSP
20	Level-2 Link State PDU	L2 LSP
24	Level-1 Complete Sequence Numbers PDU	L1 CSNP
25	Level-2 Complete Sequence Numbers PDU	L2 CSNP
26	Level-1 Partial Sequence Numbers PDU	L1 PSNP
27	Level-2 Partial Sequence Numbers PDU	L2 PSNP

Hello

Hello packets are used by routers to establish and maintain neighbor relationships. A hello packet is also an IS-to-IS hello PDU (IIH). For broadcast networks, the Level-1 routers use the Level-1 LAN IIHs; and the Level-2 routers use the Level-2 LAN IIHs. The P2P IIHs are used on point-to-point networks.

Figure 55 illustrates the hello packet format in broadcast networks, where the blue fields are the common header.

Figure 55 L1/L2 LAN IIH format

				No. of Octets
Intradomain routing protocol discriminator				1
Length indicator				1
Version/Protocol ID extension				1
ID length				1
R	R	R	PDU type	1
Version				1
Reserved				1
Maximum area address				1
Reserved/Circuit type				1
Source ID				ID length
Holding time				2
PDU length				2
R	Priority			1
LAN ID				ID length+1
Variable length fields				

Major fields of the L1/L2 LAN IIH are as follows:

- **Reserved/Circuit type**—The first six bits are reserved with a value of 0. The last two bits indicate the router type. Here, 00 means reserved, 01 indicates L1, 10 indicates L2, and 11 indicates L1/2.
- **Source ID**—System ID of the router advertising the hello packet.
- **Holding time**—If no hello packets are received from the neighbor within the holding time, the neighbor is considered down.
- **PDU length**—Total length of the PDU in bytes.
- **Priority**—DIS priority.

- **LAN ID**—Includes the system ID and a one-byte pseudonode ID.

Figure 56 shows the hello packet format on the point-to-point networks.

Figure 56 P2P IIH format

				No. of Octets
Intradomain routing protocol discriminator				1
Length indicator				1
Version/Protocol ID extension				1
ID length				1
R	R	R	PDU type	1
Version				1
Reserved				1
Maximum area address				1
Reserved/Circuit type				1
Source ID				ID length
Holding time				2
PDU length				2
Local Circuit ID				1
Variable length fields				

Instead of the priority and LAN ID fields in the LAN IIH, the P2P IIH has a Local Circuit ID field.

LSP packet format

The Link State PDU (LSP) carries link state information. LSP involves two types: Level-1 LSP and Level-2 LSP. The Level-2 LSPs are sent by the Level-2 routers, and the Level-1 LSPs are sent by the Level-1 routers. The Level-1-2 router can send both types of LSPs.

The two types of LSPs have the same format.

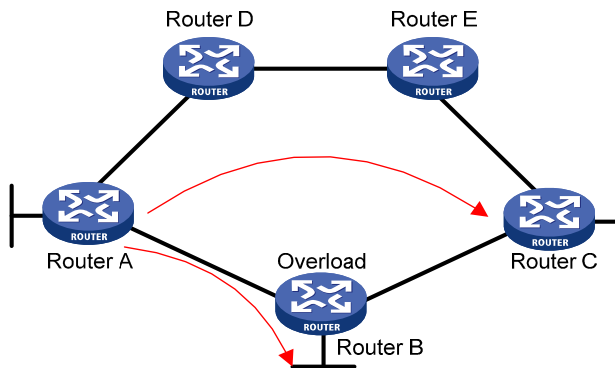
Figure 57 L1/L2 LSP format

				No. of Octets
Intradomain routing protocol discriminator				1
Length indicator				1
Version/Protocol ID extension				1
ID length				1
R	R	R	PDU type	1
Version				1
Reserved				1
Maximum area address				1
PDU length				2
Remaining lifetime				2
LSP ID				ID length+2
Sequence number				4
Checksum				2
P	ATT	OL	IS type	1
Variable length fields				

Major fields of the L1/L2 LSP are as follows:

- **PDU length**—Total length of the PDU in bytes.
- **Remaining lifetime**—LSP remaining lifetime in seconds.
- **LSP ID**—Consists of the system ID, the pseudonode ID (one byte) and the LSP fragment number (one byte).
- **Sequence number**—LSP sequence number.
- **Checksum**—LSP checksum.
- **P (Partition Repair)**—Only for L2 LSPs; it indicates whether the router supports partition repair.
- **ATT (Attachment)**—Generated by a L1/L1 router for L1 LSPs only; it indicates that the router generating the LSP is connected to multiple areas.
- **OL (LSDB Overload)**—Indicates that the LSDB is not complete because the router has run out of memory. Other routers will not send packets to the overloaded router, except packets destined to the networks directly connected to the router. For example, in [Figure 58](#), Router A forwards packets to Router C through Router B. Once other routers know the OL field of LSPs from Router B is set to 1, Router A will send packets to Router C via Router D and Router E, but still send to Router B packets destined to the network directly connected to Router B.

Figure 58 LSDB overload



- **IS type**—Type of the router generating the LSP.

SNP format

A sequence number PDU (SNP) acknowledges the latest received LSPs. It is similar to an Acknowledge packet, but more efficient.

SNP involves Complete SNP (CSNP) and Partial SNP (PSNP), which are further divided into Level-1 CSNP, Level-2 CSNP, Level-1 PSNP and Level-2 PSNP.

CSNP covers the summary of all LSPs in the LSDB to synchronize the LSDB between neighboring routers. On broadcast networks, CSNP is sent by the DIS periodically (10s by default). On point-to-point networks, CSNP is only sent during the adjacency establishment.

The CSNP packet format is shown in [Figure 59](#).

Figure 59 L1/L2 CSNP format

				No. of Octets
Intradomain routing protocol discriminator				1
Length indicator				1
Version/Protocol ID extension				1
ID length				1
R	R	R	PDU type	1
Version				1
Reserved				1
Maximum area address				1
PDU length				2
Source ID				ID length+1
Start LSP ID				ID length+2
End LSP ID				ID length+2
Variable length fields				

PSNP only contains the sequence numbers of one or multiple latest received LSPs. It can acknowledge multiple LSPs at one time. When LSDBs are not synchronized, a PSNP is used to request new LSPs from neighbors.

Figure 60 L1/L2 PSNP format

Intradomain routing protocol discriminator				No. of Octets
				1
Length indicator				1
Version/Protocol ID extension				1
ID length				1
R	R	R	PDU type	1
Version				1
Reserved				1
Maximum area address				1
PDU length				2
Source ID				ID length+1
Variable length fields				

CLV

The variable fields of PDU comprise multiple Code-Length-Value (CLV) triplets.

Figure 61 CLV format

Code	No. of Octets
Length	1
Value	Length

Table 5 shows that different PDUs contain different CLVs. Code 1 to 10 of CLV are defined in ISO 10589 (code 3 and 5 are not shown in the table), and others are defined in RFC 1195.

Table 5 CLV name and the corresponding PDU type

CLV Code	Name	PDU Type
1	Area Addresses	IIH, LSP
2	IS Neighbors (LSP)	LSP
4	Partition Designated Level2 IS	L2 LSP
6	IS Neighbors (MAC Address)	LAN IIH
7	IS Neighbors (SNPA Address)	LAN IIH
8	Padding	IIH
9	LSP Entries	SNP
10	Authentication Information	IIH, LSP, SNP
128	IP Internal Reachability Information	LSP
129	Protocols Supported	IIH, LSP
130	IP External Reachability Information	L2 LSP
131	Inter-Domain Routing Protocol Information	L2 LSP

CLV Code	Name	PDU Type
132	IP Interface Address	IIH, LSP

Supported IS-IS features

Multiple instances and processes

IS-IS supports multiple instances and processes. Multiple processes allow an IS-IS process to work in concert with a group of interfaces. A router can run multiple IS-IS processes, and each process corresponds to a unique group of interfaces.

For routers supporting VPN, each IS-IS process is associated with a VPN instance. The VPN instance is also associated with interfaces of the process.

IS-IS Graceful Restart

Graceful Restart (GR) ensures the continuity of packet forwarding when a routing protocol restarts or an active/standby switchover occurs:

- **GR Restarter**—Graceful restarting router. It must be GR capable.
- **GR Helper**—A neighbor of the GR Restarter. It helps the GR Restarter to complete the GR process.

After an IS-IS GR Restarter restarts, it must complete the following tasks to synchronize the LSDB with its neighbors:

- Obtain IS-IS neighbor information without changing adjacencies.
- Obtain the LSDB.

The GR Restarter sends an OSPF GR signal to GR Helpers so that the GR Helpers keep their adjacencies with the GR Restarter, and restores the neighbor table after receiving responses from neighbors. The GR Restarter then synchronizes the LSDB with all GR capable neighbors, calculates routes, updates its routing table and forwarding table, and removes stale routes. The IS-IS routing convergence is then complete.

IS-IS NSR

Nonstop routing (NSR) is a new feature that overcomes the application limit of GR. It backs up IS-IS link state information from the master device to the slave device. When a master/slave switchover occurs, NSR can complete link state recovery and route re-generation without requiring the cooperation of other devices.

Management tag

Management tag simplifies routing information management by carrying the management information of the IP address prefixes (to control route redistribution from other routing protocols) and BGP community and extended community attributes.

LSP fragment extension

IS-IS advertises link state information by flooding LSPs. Because one LSP carries a limited amount of link state information, IS-IS fragments LSPs. Each LSP fragment is uniquely identified by a combination of the System ID, Pseudonode ID (0 for a common LSP or a non-zero value for a Pseudonode LSP), and LSP Number (LSP fragment number) of the node or pseudo node that generated the LSP. The one-byte LSP Number field, allowing a maximum of only 256 fragments to be generated by an IS-IS router, limits the amount of link information the IS-IS router can advertise.

The LSP fragment extension feature allows an IS-IS router to generate more LSP fragments. Up to 50 additional virtual systems can be configured on the router, and each virtual system is capable of generating 256 LSP fragments to enable the IS-IS router to generate up to 13056 LSP fragments.

- Terms
 - **Originating system**—It is the router actually running IS-IS. After LSP fragment extension is enabled, additional virtual systems can be configured for the router. Originating system is the actual IS-IS process that originally runs.
 - **System ID**—System ID of the originating system
 - **Additional system ID**—Additional virtual system IDs are configured for the IS-IS router after LSP fragment extension is enabled. Each additional system ID can generate 256 LSP fragments. Both the additional system ID and the system ID must be unique in the entire routing domain.
 - **Virtual system**—A virtual system is identified by an additional system ID and generates extended LSP fragments.
 - **Original LSP**—The LSP generated by the originating system. The system ID in its LSP ID field is the system ID of the originating system.
 - **Extended LSP**—Extended LSPs are generated by virtual systems. The system ID in its LSP ID field is the virtual system ID.

After additional system IDs are configured, an IS-IS router can advertise more link state information in extended LSP fragments. Each virtual system can be considered a virtual router. An extended LSP fragment is advertised by a virtual system identified by an additional system ID.

- Operation modes:

The LSP fragment extension feature operates in the following modes:

- **Mode-1**—Applicable to a network where some routers do not support LSP fragment extension. In this mode, adjacencies are formed between the originating system and virtual systems, with the link cost from the originating system to each virtual system as 0. Each virtual system acts as a router connected to the originating system in the network, but the virtual systems are reachable through the originating system only. The IS-IS routers not supporting LSP fragment extension can operate properly without modifying the extended LSP fragments received, but some limitation is imposed on the link state information in the extended LSP fragments advertised by the virtual systems.
- **Mode-2**—Applicable to a network where all the routers support LSP fragment extension. In this mode, all the IS-IS routers know which virtual system belongs to which originating system; no limitation is imposed on the link state information of the extended LSP fragments advertised by the virtual systems.

The operation mode of LSP fragment extension is configured based on area and routing level. Mode-1 allows the routers supporting and not supporting LSP fragment extension to interoperate with each other, but it restricts the link state information in the extended fragments. Mode-2 does not restrict the link state information in the extended fragments, and is recommended for an area where all the routers are at the same routing level and support LSP fragment extension.

Dynamic host name mapping mechanism

The dynamic host name mapping mechanism provides the mappings between the host names and the system IDs for the IS-IS routers. The dynamic host name information is announced in the dynamic host name CLV of an LSP.

This mechanism also provides the mapping between a host name and the DIS of a broadcast network, which is announced in the dynamic host name TLV of a pseudonode LSP.

A host name is easier to remember than a system ID. After enabling this feature on the router, you can see the host names instead of system IDs using the **display** command.

BFD

Bidirectional forwarding detection (BFD) provides a single mechanism to quickly detect any link failures between IS-IS neighbors to reduce network convergence time. For more information about BFD, see *High Availability Configuration Guide*.

Protocols and standards

- ISO 10589 ISO IS-IS Routing Protocol
- ISO 9542 ES-IS Routing Protocol
- ISO 8348/Ad2 Network Services Access Points
- RFC 1195, *Use of OSI IS-IS for Routing in TCP/IP and Dual Environments*
- RFC 2763, *Dynamic Hostname Exchange Mechanism for IS-IS*
- RFC 2966, *Domain-wide Prefix Distribution with Two-Level IS-IS*
- RFC 3277, *IS-IS Transient Blackhole Avoidance*
- RFC 3358, *Optional Checksums in ISIS*
- RFC 3373, *Three-Way Handshake for IS-IS Point-to-Point Adjacencies*
- RFC 3567, *Intermediate System to Intermediate System (IS-IS) Cryptographic Authentication*
- RFC 3719, *Recommendations for Interoperable Networks using IS-IS*
- RFC 3786, *Extending the Number of IS-IS LSP Fragments Beyond the 256 Limit*
- RFC 3787, *Recommendations for Interoperable IP Networks using IS-IS*
- RFC 3784, *IS-IS extensions for Traffic Engineering*
- RFC 3847, *Restart signaling for IS-IS*

IS-IS configuration task list

Task	Remarks	
Configuring IS-IS basic functions	Enabling IS-IS	Required
	Configuring the IS level and circuit level	Required
	Configuring the network type of an interface as P2P	Required
	Configuring IS-IS link cost	Optional
	Specifying a priority for IS-IS	Required
Configuring IS-IS routing information control	Configuring the maximum number of ECMP routes	Optional
	Configuring IS-IS route summarization	Optional
	Advertising a default route	Optional
	Configuring IS-IS route redistribution	Optional
	Configuring IS-IS route filtering	Optional
Tuning and	Configuring IS-IS route leaking	Optional
	Specifying intervals for sending IS-IS hello and CSNP packets	Optional

Task		Remarks
optimizing IS-IS networks	Specifying the IS-IS hello multiplier	Optional
	Configuring a DIS priority for an interface	Optional
	Disabling an interface from sending or receiving IS-IS packets	Optional
	Enabling an interface to send small hello packets	Optional
	Configuring LSP parameters	Optional
	Configuring SPF parameters	Optional
	Assigning a high priority to IS-IS routes	Optional
	Setting the LSDB overload bit	Optional
	Configuring system ID to host name mappings	Optional
Enhancing IS-IS network security	Enabling the logging of neighbor state changes	Optional
	Configuring neighbor relationship authentication	Optional
	Configuring area authentication	Optional
	Configuring routing domain authentication	Optional
	Configuring IS-IS GR	Optional
	Configuring IS-IS NSR	Optional
	Configuring IS-IS FRR	Optional
	Enabling IS-IS SNMP trap	Optional
	Binding an IS-IS process with MIBs	Optional
	Configuring BFD for IS-IS	Optional

Configuring IS-IS basic functions

Configuration prerequisites

Before you configure IS-IS basic functions, complete the following tasks:

- Configure the link layer protocol.
- Configure an IP address for each interface, and make sure all neighboring nodes can reach each other at the network layer.

Enabling IS-IS

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable the IS-IS routing process and enter its view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	Not enabled by default.
3. Assign a network entity title (NET).	network-entity <i>net</i>	Not assigned by default.
4. Return to system view.	quit	N/A

Step	Command	Remarks
5. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
6. Enable an IS-IS process on the interface.	isis enable [<i>process-id</i>]	Disabled by default.

Configuring the IS level and circuit level

If only one area is available, HP recommends you to perform the following operations:

- Configure the IS level of all routers as Level-1 or Level-2 rather than different levels because the routers do not need to maintain two identical LSDBs.
- Configure the IS level as Level-2 on all routers in an IP network for scalability.

For an interface of a Level-1 (or Level-2) router, the circuit level can only be Level-1 (or Level-2). For an interface of a Level-1-2 router, the default circuit level is Level-1-2; if the router only needs to form Level-1 (or Level-2) neighbor relationships, then configure the circuit level for its interfaces as Level-1 (or Level-2) to limit neighbor relationship establishment.

To configure the IS level and circuit level:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Specify the IS level.	is-level { level-1 level-1-2 level-2 }	Optional. The default is Level-1-2.
4. Return to system view.	quit	N/A
5. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
6. Specify the circuit level.	isis circuit-level [level-1 level-1-2 level-2]	Optional. By default, an interface can establish the Level-1 or Level-2 adjacency.

Configuring the network type of an interface as P2P

Perform this task only for a broadcast network that has up to two attached routers.

Interfaces with different network types operate differently. For example, broadcast interfaces on a network must elect the DIS and flood CSNP packets to synchronize the LSDBs, but P2P interfaces on a network do not need to elect the DIS, and have a different LSDB synchronization mechanism.

If only two routers exist on a broadcast network, configure the network type of attached interfaces as P2P to avoid DIS election and CSNP flooding, saving network bandwidth and speeding up network convergence.

To configure the network type of an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the network type for the interface as P2P.	isis circuit-type p2p	Optional. By default, the network type of an interface depends on the physical media. The network type of a VLAN interface is broadcast.

Configuring IS-IS routing information control

Configuration prerequisites

Before configuring IS-IS routing information control, complete the following tasks:

- Configure network layer addresses for interfaces, and make sure that adjacent nodes can reach each other at the network layer.
- Enable IS-IS.

Configuring IS-IS link cost

The IS-IS cost of an interface is determined in the following order:

1. IS-IS cost specified in interface view.
2. IS-IS cost specified in system view. The cost is applied to the interfaces associated with the IS-IS process.
3. Automatically calculated cost. If the cost style is **wide** or **wide-compatible**, IS-IS automatically calculates the cost using the formula: $\text{interface cost} = (\text{bandwidth reference value} / \text{interface bandwidth}) \times 10$, which is in the range of 1 to 16777214. For other cost styles, [Table 6](#) applies.

Table 6 Automatic cost calculation scheme for cost styles other than wide and wide-compatible

Interface bandwidth	Interface cost
≤ 10 Mbps	60
≤ 100 Mbps	50
≤ 155 Mbps	40
≤ 622 Mbps	30
≤ 2500 Mbps	20
> 2500 Mbps	10

4. If none of the above costs are used, a default cost of 10 applies.

Configuring an IS-IS cost for an interface

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Specify an IS-IS cost style.	cost-style { narrow wide wide-compatible { compatible narrow-compatible } [relax-spf-limit] }	Optional. narrow by default.
4. Return to system view.	quit	N/A
5. Enter interface view.	interface <i>interface-type interface-number</i>	N/A
6. Specify a cost for the interface.	isis cost <i>value</i> [level-1 level-2]	Optional. No cost is specified for the interface by default.

Configuring a global IS-IS cost

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Specify an IS-IS cost style.	cost-style { narrow wide wide-compatible { compatible narrow-compatible } [relax-spf-limit] }	Optional. narrow by default.
4. Specify a global IS-IS cost.	circuit-cost <i>value</i> [level-1 level-2]	By default, no global cost is specified.

Enabling automatic IS-IS cost calculation

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Specify an IS-IS cost style.	cost-style { wide wide-compatible }	narrow by default.
4. Enable automatic IS-IS cost calculation.	auto-cost enable	Disabled by default.
5. Configure a bandwidth reference value for automatic IS-IS cost calculation.	bandwidth-reference <i>value</i>	Optional. 100 Mbps by default.

Specifying a priority for IS-IS

A router can run multiple routing protocols. When routes to the same destination are found by multiple routing protocols, the route learned by the protocol with the highest priority can be adopted. You can reference a routing policy to specify a priority for specific routes. For information about routing policy, see "[Configuring routing policies.](#)"

To configure the priority of IS-IS:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Specify a priority for IS-IS.	preference { route-policy <i>route-policy-name</i> <i>preference</i> } *	15 by default.

Configuring the maximum number of ECMP routes

Perform this task to implement load sharing over ECMP routes.

To configure the maximum number of ECMP routes:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Specify the maximum number of ECMP routes for load balancing.	maximum load-balancing <i>number</i>	8 by default.

Configuring IS-IS route summarization

This task allows you to configure a summary route so routes falling into the network range of the summary route are summarized into one route for advertisement. Doing so can reduce the size of routing tables, as well as the scale of LSP and LSDB. Both IS-IS routes and redistributed routes can be summarized.

The router summarizes only the routes in the locally generated LSPs. The cost of the summary route is the lowest one among the costs of summarized routes.

To configure route summarization:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure IS-IS route summarization.	summary <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } [avoid-feedback generate_null0_route tag <i>tag</i> [level-1 level-1-2 level-2]] *	By default, no route summarization is configured.

Advertising a default route

A router running IS-IS cannot redistribute any default routes or advertise a default route to neighbors. Use the following commands to advertise a default route of 0.0.0.0/0 to the same level neighbors.

The default route is only advertised to routers at the same level. You can use a routing policy to generate the default route only when a local routing entry is matched by the policy.

To advertise a default route:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Advertise a default route.	default-route-advertise [route-policy <i>route-policy-name</i> [level-1 level-1-2 level-2]] *	By default, the function is disabled.

Configuring IS-IS route redistribution

Redistributing large numbers of routes on a device may affect the performance of other devices in the network. If this happens, you can configure a limit on the number of redistributed routes in order to limit the number of routes to be advertised.

Only active routes can be redistributed. Use the **display ip routing-table protocol** command to display route state information.

To configure IS-IS route redistribution from other routing protocols:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Redistribute routes from another routing protocol.	import-route <i>protocol</i> [<i>process-id</i> all-processes allow-ibgp] [cost <i>cost</i> cost-type { external internal }] [level-1 level-1-2 level-2] route-policy <i>route-policy-name</i> tag tag] *	By default: <ul style="list-style-type: none"> No route is redistributed. If no level is specified, routes are redistributed into the Level-2 routing table.
4. Configure the maximum number of redistributed Level 1/Level 2 IPv4 routes.	import-route limit <i>number</i>	Optional. The default varies with devices.

Configuring IS-IS route filtering

You can reference a configured ACL, IP prefix list, or routing policy to filter routes calculated from the received LSPs and the routes redistributed from other routing protocols.

Filtering routes calculated from received LSPs

IS-IS saves the LSPs received from neighbors in the LSDB, uses the SPF algorithm to calculate the shortest path tree with itself as the root, and installs the routes into the IS-IS routing table.

By referencing a configured ACL, IP prefix list, or routing policy, you can filter the calculated routes. Only the routes matching the filter can be added into the IS-IS routing table.

To filter routes calculated from received LSPs:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Filter routes calculated from received LSPs.	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> route-policy <i>route-policy-name</i> } import	By default, no filtering is configured.

Filtering redistributed routes

IS-IS can redistribute routes from other routing protocols (or other IS-IS processes), add them into the IS-IS routing table, and advertise them in LSPs.

By reference a configured ACL, IP prefix list, or routing policy, you can filter redistributed routes and only the routes matching the filter can be added into the IS-IS routing table and advertised to neighbors.

To configure the filtering of redistributed routes:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure the filtering of routes redistributed from another routing protocol or IS-IS process.	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> route-policy <i>route-policy-name</i> } export [<i>protocol</i> [<i>process-id</i>]]	Not configured by default.

Configuring IS-IS route leaking

With IS-IS route leaking enabled, the Level-1-2 router can advertise the routing information of other Level-1 areas and Level-2 area routing information to Level-1 routers.

If a filter policy is specified, only routes passing it can be advertised into Level-1 area.

You can specify a routing policy in the **import-route isis level-2 into level-1** command to filter routes from Level-2 to Level-1. Other routing policies specified for route reception and redistribution does not affect the route leaking.

To configure IS-IS route leaking:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Enable IS-IS route leaking.	import-route isis level-2 into level-1 [filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> route-policy <i>route-policy-name</i> } tag <i>tag</i>] *	Disabled by default.

Tuning and optimizing IS-IS networks

Configuration prerequisites

Before you tune and optimize IS-IS networks, complete the following tasks:

- Configure IP addresses for interfaces, and make adjacent nodes can reach each other at the network layer.
- Enable IS-IS.

Specifying intervals for sending IS-IS hello and CSNP packets

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Specify the interval for sending hello packets.	isis timer hello <i>seconds</i> [level-1 level-2]	Optional. 10 seconds by default.
4. Specify the interval for sending CSNP packets on the DIS of a broadcast network.	isis timer csnp <i>seconds</i> [level-1 level-2]	Optional. 10 seconds by default.

NOTE:

The interval between hello packets sent by the DIS is 1/3 the hello interval set with the **isis timer hello** command.

Specifying the IS-IS hello multiplier

If a neighbor receives no hello packets from the router within the advertised hold time, it considers the router down and recalculates the routes. The hold time is the hello multiplier multiplied by the hello interval.

On a broadcast link, Level-1 and Level-2 hello packets are advertised separately; therefore, you need to set a hello multiplier for each level.

On a P2P link, Level-1 and Level-2 hello packets are advertised in P2P hello packets, so you do not need to specify Level-1 or Level-2.

To specify the IS-IS hello multiplier:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Specify the number of hello packets a neighbor must miss before declaring the router is down.	isis timer holding-multiplier <i>value</i> [level-1 level-2]	Optional. 3 by default.

Configuring a DIS priority for an interface

On an IS-IS broadcast network, you must elect a router as the DIS at a routing level. You can specify a DIS priority at a level for an interface. The greater the interface's priority, the more likely it becomes the DIS. If multiple routers in the broadcast network have the same highest DIS priority, the router with the highest MAC address becomes the DIS.

To specify a DIS priority for an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Specify a DIS priority for the interface.	isis dis-priority <i>value</i> [level-1 level-2]	Optional. 64 by default.

Disabling an interface from sending or receiving IS-IS packets

After being disabled from sending or receiving hello packets, an interface cannot form a neighbor relationship, but can advertise directly connected networks in LSPs through other interfaces. This can save bandwidth and CPU resources, and ensures other routers know networks directly connected to the interface.

To disable an interface from sending and receiving IS-IS packets:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Disable the interface from sending and receiving IS-IS packets.	isis silent	Not disabled by default.

Enabling an interface to send small hello packets

IS-IS messages cannot be fragmented at the IP layer because they are directly encapsulated into frames. Any two IS-IS neighboring routers must negotiate a common MTU. To avoid sending big hellos for saving bandwidth, enable the interface to send small hello packets without CLVs.

To enable an interface to send small hello packets:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable the interface to send small hello packets without CLVs.	isis small-hello	Standard hello packets are sent by default.

Configuring LSP parameters

Configuring LSP timers

- Specify the maximum age of LSPs.

Each LSP has an age that decreases in the LSDB. Any LSP with an age of 0 is deleted from the LSDB. You can adjust the age value based on the scale of a network.

To specify the maximum age of LSPs:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Specify the maximum LSP age.	timer lsp-max-age <i>seconds</i>	Optional. 1200 seconds by default.

- Specify the LSP refresh interval and generation interval.

Each router needs to refresh LSPs generated by itself at a configurable interval and send them to other routers to prevent valid routes from being aged out. A smaller refresh interval speeds up network convergence but consumes more bandwidth.

When the network topology changes, for example, a neighbor is down or up, or the interface metric, system ID, or area ID is changed, the router generates an LSP after a configurable interval. If such a change occurs frequently, excessive LSPs are generated, consuming a large amount of router resources and bandwidth. To solve the problem, you can adjust the LSP generation interval.

To specify the LSP refresh interval and generation interval:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Specify the LSP refresh interval.	timer lsp-refresh <i>seconds</i>	Optional. 900 seconds by default.
4. Specify the LSP generation interval.	timer lsp-generation <i>maximum-interval</i> [<i>initial-interval</i> [<i>second-wait-interval</i>]] [level-1 level-2]	Optional. 2 seconds by default.

- Specify LSP sending intervals.

If a change occurs in the LSDB, IS-IS advertises the changed LSP to neighbors. You can specify the minimum interval for sending such LSPs.

On a P2P link, IS-IS requires an advertised LSP be acknowledged. If no acknowledgement is received within a configurable interval, IS-IS will retransmit the LSP.

To configure LSP sending intervals:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Specify the minimum interval for sending LSPs and the maximum LSP number that can be sent at a time.	isis timer lsp <i>time</i> [count <i>count</i>]	Optional. By default, the minimum interval is 33 milliseconds, and the maximum LSP number that can be sent at a time is 5.
4. Specify the LSP retransmission interval on a P2P link.	isis timer retransmit <i>seconds</i>	Optional. 5 seconds by default. Configure a proper LSP retransmission interval to avoid unnecessary retransmissions.

Specifying LSP lengths

IS-IS messages cannot be fragmented at the IP layer because they are directly encapsulated in frames. IS-IS routers in an area must send LSPs smaller than the smallest interface MTU in this area.

If the IS-IS routers have different interface MTUs, HP recommends configuring the maximum size of generated LSP packets to be smaller than the smallest interface MTU in this area. If they are not, the routers must dynamically adjust the LSP packet size to fit the smallest interface MTU, which takes time and affects other services.

To specify LSP lengths:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Specify the maximum length of generated Level-1 LSPs or Level-2 LSPs.	lsp-length originate <i>size</i> [level-1 level-2]	1497 bytes by default.
4. Specify the maximum length of received LSPs.	lsp-length receive <i>size</i>	1497 bytes by default.

Enabling LSP flash flooding

Changed LSPs may trigger SPF recalculation, so you can enable LSP flash flooding to advertise the changed LSPs before the router recalculates routes. Doing so can speed up network convergence.

To enable LSP flash flooding:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Enable LSP flash flooding.	flash-flood [flood-count <i>flooding-count</i> max-timer-interval <i>flooding-interval</i> [level-1 level-2]] *	Not enabled by default.

Enabling LSP fragment extension

After LSP fragment extension is enabled for an IS-IS process, the MTUs of all the interfaces running the IS-IS process must not be less than 512; otherwise, LSP fragment extension will not take effect.

At least one virtual system must be configured for the router to generate extended LSP fragments. An IS-IS process allows 50 virtual systems.

To enable LSP fragment extension:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Enable LSP fragment extension and specify the working mode.	lsp-fragments-extend [[level-1 level-1-2 level-2] [mode-1 mode-2]] *	Not enabled by default.
4. Configure a virtual system ID.	virtual-system <i>virtual-system-id</i>	Not configured by default.

Configuring SPF parameters

When the LSDB changes on a router, a route calculation starts. Frequent route calculations consume a lot of system resources. You can set an appropriate interval for SPF calculations to improve efficiency.

To configure the SPF parameters:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure the SPF calculation interval.	timer spf <i>maximum-interval</i> [<i>initial-interval</i> [<i>second-wait-interval</i>]]	Optional. The default SPF calculation interval is 10 seconds.

Assigning a high priority to IS-IS routes

An IS-IS topology change causes network convergence. By assigning a high priority to specific IS-IS routes, you can achieve faster network convergence.

To assign a high priority to IS-IS routes:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A

Step	Command	Remarks
3.	Assign a high priority to IS-IS routes.	Optional. Not assigned by default. If no IS-IS route is assigned a high priority, IS-IS host routes are processed first in network convergence because they have higher priority than other types of IS-IS routes.

Setting the LSDB overload bit

By setting the overload bit in sent LSPs, a router informs other routers of a failure that makes it incapable of routing and forwarding packets.

When an IS-IS router cannot record the complete LSDB due to running out of memory or some other reasons, it will calculate wrong routes. To make troubleshooting easier, you can temporarily isolate the router from the IS-IS network by setting the overload bit.

To set the LSDB overload bit:

Step	Command	Remarks
1.	Enter system view.	N/A
2.	Enter IS-IS view.	N/A
3.	Set the overload bit.	Not set by default.

Configuring system ID to host name mappings

In IS-IS, a system ID identifies a router or host uniquely. A system ID has a fixed length of 6 bytes. When an administrator needs to view IS-IS neighbor information, routing table or LSDB information, using the system IDs in dotted decimal notation is not convenient. To solve it, you can configure the mappings between system IDs and host names, as host names are easier to remember and use.

Such mappings can be configured manually or dynamically. Note the following:

- Using the **display isis lsdb** command on a router configured with dynamic system ID to host name mapping displays router names rather than system IDs.
- If you configure both dynamic and static system ID to host name mappings on a router, the host name for dynamic system ID to host name mapping applies.

Configuring a static system ID to host name mapping

Step	Command	Remarks
1.	Enter system view.	N/A
2.	Enter IS-IS view.	N/A

Step	Command	Remarks
3.	Configure a system ID to host name mapping for a remote IS. is-name map <i>sys-id map-sys-name</i>	A system ID can only correspond to a host name.

Configuring dynamic system ID to host name mapping

Configure a static system ID to host name mapping for any other router in a network. When a new router is added into the network or a mapping must be modified, perform configuration on all routers.

You can configure dynamic system ID to host name mapping. To do so, you must configure a host name for each router in the network. Each router advertises the host name in dynamic host name CLVs to other routers. Then, all routers in the network have all the mappings to generate a mapping table.

In addition, you can configure a name for the DIS in a broadcast network to help check the origin of LSPs in the LSDB.

To configure dynamic system ID to host name mapping:

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Enter IS-IS view. isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3.	Specify a host name for the router. is-name <i>sys-name</i>	Not specified by default.
4.	Return to system view. quit	N/A
5.	Enter interface view. interface <i>interface-type</i> <i>interface-number</i>	N/A
6.	Configure a DIS name. isis dis-name <i>symbolic-name</i>	Optional. Not configured by default. This command takes effect only on a router with dynamic system ID to host name mapping configured. This command is not supported on P2P interfaces.

Enabling the logging of neighbor state changes

Logging of neighbor state changes enables the router to output neighbor state changes to the console terminal.

To enable the logging of neighbor state changes:

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Enter IS-IS view. isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3.	Enable the logging of neighbor state changes. log-peer-change	Enabled by default.

Enhancing IS-IS network security

To enhance the security of an IS-IS network, you can configure IS-IS authentication. IS-IS authentication involves neighbor relationship authentication, area authentication and routing domain authentication.

Configuration prerequisites

Before you enhance IS-IS network security, complete the following tasks:

- Configure IP addresses for interfaces, and make sure that all neighboring nodes can reach each other at the network layer.
- Enable IS-IS.

Configuring neighbor relationship authentication

With neighbor relationship authentication configured, an interface adds the password in the specified mode into hello packets to the peer and checks the password in the received hello packets. If the authentication succeeds, it forms the neighbor relationship with the peer.

Follow these guidelines when you configure neighbor relationship authentication:

- The authentication mode and password at both ends must be identical.
- The **level-1** and **level-2** keywords are configurable on an interface that has IS-IS enabled.
- If you configure an authentication mode and a password without specifying a level, the authentication mode and password apply to both Level-1 and Level-2.
- If neither **ip** nor **osi** is specified, the OSI related fields in LSPs are checked.

To configure neighbor relationship authentication:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Specify the authentication mode and password.	isis authentication-mode { md5 simple } [cipher] <i>password</i> [level-1 level-2] [ip osi]	By default, no authentication is configured.

Configuring area authentication

Area authentication enables a router not to install routing information from untrusted routers into the Level-1 LSDB. The router encapsulates the authentication password in the specified mode into Level-1 packets (LSP, CSNP, and PSNP) and checks the password in received Level-1 packets.

Routers in a common area must have the same authentication mode and password.

To configure area authentication:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A

Step	Command	Remarks
3. Specify the area authentication mode and password.	area-authentication-mode { md5 simple } [cipher] <i>password</i> [ip osi]	By default, no area authentication is configured.

Configuring routing domain authentication

Routing domain authentication prevents untrusted routing information from entering into a routing domain. A router with the authentication configured encapsulates the password in the specified mode into Level-2 packets (LSP, CSNP, PSNP) and check the password in received Level-2 packets.

All the routers in the backbone must have the same authentication mode and password.

To configure routing domain authentication:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Specify the routing domain authentication mode and password.	domain-authentication-mode { md5 simple } [cipher] <i>password</i> [ip osi]	By default, no routing domain authentication is configured.

Configuring IS-IS GR

Restarting IS-IS on a router causes network disconnections and route reconvergence.

With the Graceful Restart (GR) feature, the restarting router—known as the GR Restarter—can notify the event to its GR capable neighbors. GR capable neighbors—known as the GR Helpers—will keep their adjacencies with the router within a configurable GR interval. After the restart, the router contacts its neighbors to retrieve its routing table.

During this process, the network keeps stable.

The IS-IS GR and IS-IS NSR features are mutually exclusive.

To configure GR on the GR Restarter and GR Helper:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IS-IS, and enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	Disabled by default.
3. Enable the GR capability for IS-IS.	graceful-restart	Disabled by default.
4. Set the Graceful Restart interval.	graceful-restart interval <i>timer</i>	300 seconds by default. The Graceful Restart interval is set as the holding time in the hello PDUs. Within the interval, the neighbors will keep their adjacency with the GR Restarter.

Step	Command	Remarks
5. Suppress the SA bit during restart.	graceful-restart suppress-sa	Optional. By default, the SA bit is not suppressed. By enabling the GR Restarter to suppress the Suppress-Advertisement (SA) bit in the hello PDUs, the neighbors will still advertise their adjacency with the GR Restarter.

Configuring IS-IS NSR

According to the GR feature, after a master/slave switchover, the GR Restarter obtains routing information from its neighbors and the IS-IS process on the new master needs to learn all routes. If the network topology has changed during the switchover period, removed routes cannot be updated to the device, which may cause black hole routes.

NSR is introduced to solve the problem. It backs up IS-IS link state information from the master device to the slave device. After a master/slave switchover, NSR can complete link state recovery and route re-generation without requiring the cooperation of other devices.

The IS-IS NSR and IS-IS GR features are mutually exclusive.

To configure IS-IS NSR:

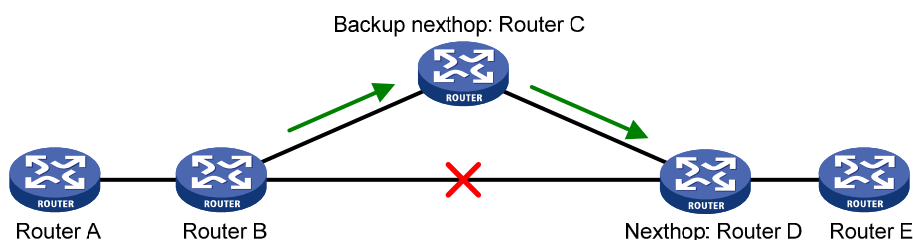
Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [process-id] [vpn-instance vpn-instance-name]	N/A
3. Enable IS-IS NSR.	non-stop-routing	Disabled by default.
4. Set the NSR interval.	non-stop-routing interval interval-value	0 seconds by default, that is, no NSR interval is configured.

Configuring IS-IS FRR

When a link fails, the packets on the path are discarded, or a routing loop occurs until IS-IS completes the routing convergence based on the new network topology.

You can enable IS-IS fast reroute (FRR) to reduce traffic recovery time.

Figure 62 Network diagram for IS-IS FRR



In [Figure 62](#), after you enable FRR on Router B, IS-IS automatically calculates or designates a backup next hop when a link failure is detected. In this way, packets are directed to the backup next hop to reduce

traffic recovery time. Meanwhile, IS-IS calculates the shortest path based on the new network topology, and forwards packets over the path after network convergence.

You can either enable IS-IS FRR to calculate a backup next hop automatically, or to designate a backup next hop with a routing policy for routes matching specific criteria.

Configuration prerequisites

Before you configure IS-IS FRR, complete the following tasks:

- Configure IP addresses for interfaces, and make sure that all neighboring nodes can reach each other at the network layer.
- Enable IS-IS.

Configuration guidelines

- Do not use FRR and BFD at the same time. Otherwise, FRR may fail to take effect.
- The automatic backup next hop calculation of FRR and that of TE are mutually exclusive.

Configuring IS-IS FRR to automatically calculate a backup next hop

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the source address of echo packets.	bfd echo-source-ip <i>ip-address</i>	Not configured by default
3. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
4. Enable IS-IS FRR to automatically calculate a backup next hop.	fast-reroute auto	Not configured by default

Configuring IS-IS FRR to designate a backup next hop with a routing policy

You can use the **apply fast-reroute backup-interface** command to specify a backup next hop in a routing policy for routes matching specific criteria. For more information about the **apply fast-reroute backup-interface** command and routing policy configurations, see the "[Configuring routing policies.](#)"

To configure IS-IS FRR:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the source address of echo packets.	bfd echo-source-ip <i>ip-address</i>	Not configured by default.
3. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
4. Enable IS-IS FRR to designate a backup next hop by using a routing policy.	fast-reroute route-policy <i>route-policy-name</i>	Not configured by default.

Enabling IS-IS SNMP trap

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Enable SNMP trap.	is-snmp-traps enable	Enabled by default

Binding an IS-IS process with MIBs

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Bind the IS-IS process with MIBs.	isis mib-binding <i>process-id</i>	By default, MIBs are bound with IS-IS process 1.

Configuring BFD for IS-IS

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type interface-number</i>	N/A
3. Enable IS-IS on the interface.	isis enable [<i>process-id</i>]	Disabled by default.
4. Enable BFD on the IS-IS interface.	isis bfd enable	Not enabled by default.

Displaying and maintaining IS-IS

Task	Command	Remarks
Display brief IS-IS configuration information.	display isis brief [<i>process-id</i> vpn-instance <i>vpn-instance-name</i>] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the status of IS-IS debug switches.	display isis debug-switches { <i>process-id</i> vpn-instance <i>vpn-instance-name</i> } [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the IS-IS Graceful Restart state.	display isis graceful-restart status [level-1 level-2] [<i>process-id</i> vpn-instance <i>vpn-instance-name</i>] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display information about IS-IS enabled interfaces.	display isis interface [statistics [<i>interface-type interface-number</i>] [verbose]] [<i>process-id</i> vpn-instance <i>vpn-instance-name</i>] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display IS-IS LSDB information.	display isis lsdb [[I1 I2 level-1 level-2] [lsp-id <i>lspid</i> lsp-name <i>lspname</i>] local verbose] * [<i>process-id</i> vpn-instance <i>vpn-instance-name</i>] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display the host-name-to-system-ID mapping table.	display isis name-table [<i>process-id</i> vpn-instance <i>vpn-instance-name</i>] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display IS-IS neighbor information.	display isis peer [statistics verbose] [<i>process-id</i> vpn-instance <i>vpn-instance-name</i>] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display IS-IS IPv4 routing information.	display isis route [ipv4] [[level-1 level-2] verbose] * [<i>process-id</i> vpn-instance <i>vpn-instance-name</i>] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display IS-IS SPF calculation log information.	display isis spf-log [<i>process-id</i> vpn-instance <i>vpn-instance-name</i>] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display IS-IS statistics.	display isis statistics [level-1 level-1-2 level-2] [<i>process-id</i> vpn-instance <i>vpn-instance-name</i>] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Clear IS-IS process data structure information.	reset isis all [<i>process-id</i> vpn-instance <i>vpn-instance-name</i>]	Available in user view
Clear the data structure information of an IS-IS neighbor.	reset isis peer <i>system-id</i> [<i>process-id</i> vpn-instance <i>vpn-instance-name</i>]	Available in user view

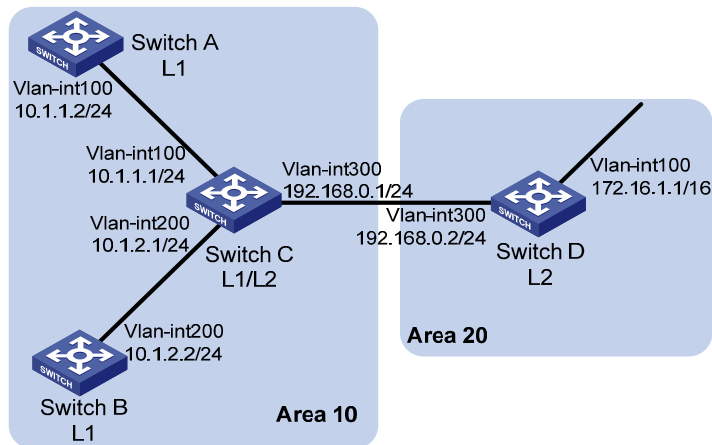
IS-IS configuration examples

IS-IS basic configuration

Network requirements

As shown in [Figure 63](#), Switch A, B, C, and D reside in an IS-IS AS. Switch A and B are Level-1 switches, Switch D is a Level-2 switch, and Switch C is a Level-1-2 switch. Switch A, B, and C are in Area 10, and Switch D is in Area 20.

Figure 63 Network diagram for IS-IS basic configuration



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure IS-IS:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] isis 1
[SwitchA-isis-1] is-level level-1
[SwitchA-isis-1] network-entity 10.0000.0000.0001.00
[SwitchA-isis-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] isis enable 1
[SwitchA-Vlan-interface100] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] isis 1
[SwitchB-isis-1] is-level level-1
[SwitchB-isis-1] network-entity 10.0000.0000.0002.00
[SwitchB-isis-1] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] isis enable 1
[SwitchB-Vlan-interface200] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] isis 1
[SwitchC-isis-1] network-entity 10.0000.0000.0003.00
[SwitchC-isis-1] quit
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] isis enable 1
[SwitchC-Vlan-interface100] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] isis enable 1
[SwitchC-Vlan-interface200] quit
```

```
[SwitchC] interface vlan-interface 300
[SwitchC-Vlan-interface300] isis enable 1
[SwitchC-Vlan-interface300] quit
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] isis 1
[SwitchD-isis-1] is-level level-2
[SwitchD-isis-1] network-entity 20.0000.0000.0004.00
[SwitchD-isis-1] quit
[SwitchD] interface vlan-interface 100
[SwitchD-Vlan-interface100] isis enable 1
[SwitchD-Vlan-interface100] quit
[SwitchD] interface vlan-interface 300
[SwitchD-Vlan-interface300] isis enable 1
[SwitchD-Vlan-interface300] quit
```

3. Verify the configuration:

Display the IS-IS LSDB of each switch to check the LSP integrity.

```
[SwitchA] display isis lsdb
```

```
Database information for ISIS(1)
```

```
-----
```

Level-1 Link State Database

LSPID	Seq Num	Checksum	Holdtime	Length	ATT/P/OL
0000.0000.0001.00-00*	0x00000004	0xdf5e	1096	68	0/0/0
0000.0000.0002.00-00	0x00000004	0xee4d	1102	68	0/0/0
0000.0000.0002.01-00	0x00000001	0xdaaf	1102	55	0/0/0
0000.0000.0003.00-00	0x00000009	0xcaa3	1161	111	1/0/0
0000.0000.0003.01-00	0x00000001	0xadda	1112	55	0/0/0

*-Self LSP, +-Self LSP(Extended), ATT-Attached, P-Partition, OL-Overload

```
[SwitchB] display isis lsdb
```

```
Database information for ISIS(1)
```

```
-----
```

Level-1 Link State Database

LSPID	Seq Num	Checksum	Holdtime	Length	ATT/P/OL
0000.0000.0001.00-00	0x00000006	0xdb60	988	68	0/0/0
0000.0000.0002.00-00*	0x00000008	0xe651	1189	68	0/0/0
0000.0000.0002.01-00*	0x00000005	0xd2b3	1188	55	0/0/0
0000.0000.0003.00-00	0x00000014	0x194a	1190	111	1/0/0
0000.0000.0003.01-00	0x00000002	0xabdb	995	55	0/0/0

*-Self LSP, +-Self LSP(Extended), ATT-Attached, P-Partition, OL-Overload
[SwitchC] display isis lsdb

Database information for ISIS(1)

Level-1 Link State Database

LSPID	Seq Num	Checksum	Holdtime	Length	ATT/P/OL
0000.0000.0001.00-00	0x00000006	0xdb60	847	68	0/0/0
0000.0000.0002.00-00	0x00000008	0xe651	1053	68	0/0/0
0000.0000.0002.01-00	0x00000005	0xd2b3	1052	55	0/0/0
0000.0000.0003.00-00*	0x00000014	0x194a	1051	111	1/0/0
0000.0000.0003.01-00*	0x00000002	0xabdb	854	55	0/0/0

*-Self LSP, +-Self LSP(Extended), ATT-Attached, P-Partition, OL-Overload

Level-2 Link State Database

LSPID	Seq Num	Checksum	Holdtime	Length	ATT/P/OL
0000.0000.0003.00-00*	0x00000012	0xc93c	842	100	0/0/0
0000.0000.0004.00-00	0x00000026	0x331	1173	84	0/0/0
0000.0000.0004.01-00	0x00000001	0xee95	668	55	0/0/0

*-Self LSP, +-Self LSP(Extended), ATT-Attached, P-Partition, OL-Overload
[SwitchD] display isis lsdb

Database information for ISIS(1)

Level-2 Link State Database

LSPID	Seq Num	Checksum	Holdtime	Length	ATT/P/OL
0000.0000.0003.00-00	0x00000013	0xc73d	1003	100	0/0/0
0000.0000.0004.00-00*	0x0000003c	0xd647	1194	84	0/0/0
0000.0000.0004.01-00*	0x00000002	0xec96	1007	55	0/0/0

*-Self LSP, +-Self LSP(Extended), ATT-Attached, P-Partition, OL-Overload

Display the IS-IS routing information of each switch. Level-1 switches must have a default route with the next hop being the Level-1-2 switch. The Level-2 switch must have both routing information of Level-1 and Level-2.

[SwitchA] display isis route

Route information for ISIS(1)

 ISIS(1) IPv4 Level-1 Forwarding Table

IPv4 Destination	IntCost	ExtCost	ExitInterface	NextHop	Flags
10.1.1.0/24	10	NULL	Vlan100	Direct	D/L/-
10.1.2.0/24	20	NULL	Vlan100	10.1.1.1	R/-/-
192.168.0.0/24	20	NULL	Vlan100	10.1.1.1	R/-/-
0.0.0.0/0	10	NULL	Vlan100	10.1.1.1	R/-/-

Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set
 [SwitchC] display isis route

Route information for ISIS(1)

ISIS(1) IPv4 Level-1 Forwarding Table

IPv4 Destination	IntCost	ExtCost	ExitInterface	NextHop	Flags
192.168.0.0/24	10	NULL	Vlan300	Direct	D/L/-
10.1.1.0/24	10	NULL	Vlan100	Direct	D/L/-
10.1.2.0/24	10	NULL	Vlan200	Direct	D/L/-

Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set

ISIS(1) IPv4 Level-2 Forwarding Table

IPv4 Destination	IntCost	ExtCost	ExitInterface	NextHop	Flags
192.168.0.0/24	10	NULL	Vlan300	Direct	D/L/-
10.1.1.0/24	10	NULL	Vlan100	Direct	D/L/-
10.1.2.0/24	10	NULL	Vlan200	Direct	D/L/-
172.16.0.0/16	20	NULL	Vlan300	192.168.0.2	R/-/-

Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set
 [SwitchD] display isis route

Route information for ISIS(1)

ISIS(1) IPv4 Level-2 Forwarding Table

IPv4 Destination	IntCost	ExtCost	ExitInterface	NextHop	Flags
192.168.0.0/24	10	NULL	Vlan300	Direct	D/L/-
10.1.1.0/24	20	NULL	Vlan300	192.168.0.1	R/-/-
10.1.2.0/24	20	NULL	Vlan300	192.168.0.1	R/-/-
172.16.0.0/16	10	NULL	Vlan100	Direct	D/L/-

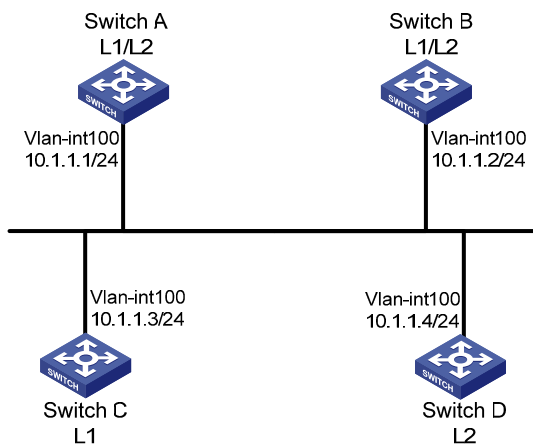
Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set

DIS election configuration

Network requirements

As shown in [Figure 64](#), Switch A, B, C, and D reside in IS-IS area 10 on a broadcast network (Ethernet). Switch A and Switch B are Level-1-2 switches, Switch C is a Level-1 switch, and Switch D is a Level-2 switch. Change the DIS priority of Switch A to make it elected as the Level-1-2 DIS router.

Figure 64 Network diagram for DIS selection



Configuration procedure

1. Configure an IP address for each interface. (Details not shown.)
2. Enable IS-IS:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] isis 1
[SwitchA-isis-1] network-entity 10.0000.0000.0001.00
[SwitchA-isis-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] isis enable 1
[SwitchA-Vlan-interface100] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] isis 1
[SwitchB-isis-1] network-entity 10.0000.0000.0002.00
[SwitchB-isis-1] quit
```



```

[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] isis enable 1
[SwitchB-Vlan-interface100] quit
# Configure Switch C.
<SwitchC> system-view
[SwitchC] isis 1
[SwitchC-isis-1] network-entity 10.0000.0000.0003.00
[SwitchC-isis-1] is-level level-1
[SwitchC-isis-1] quit
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] isis enable 1
[SwitchC-Vlan-interface100] quit
# Configure Switch D.
<SwitchD> system-view
[SwitchD] isis 1
[SwitchD-isis-1] network-entity 10.0000.0000.0004.00
[SwitchD-isis-1] is-level level-2
[SwitchD-isis-1] quit
[SwitchD] interface vlan-interface 100
[SwitchD-Vlan-interface100] isis enable 1
[SwitchD-Vlan-interface100] quit
# Display information about IS-IS neighbors of Switch A.
[SwitchA] display isis peer

```

Peer information for ISIS(1)

```

-----
System Id: 0000.0000.0002
Interface: Vlan-interface100      Circuit Id: 0000.0000.0003.01
State: Up      HoldTime: 21s      Type: L1(L1L2)      PRI: 64

System Id: 0000.0000.0003
Interface: Vlan-interface100      Circuit Id: 0000.0000.0003.01
State: Up      HoldTime: 27s      Type: L1              PRI: 64

System Id: 0000.0000.0002
Interface: Vlan-interface100      Circuit Id: 0000.0000.0004.01
State: Up      HoldTime: 28s      Type: L2(L1L2)      PRI: 64

System Id: 0000.0000.0004
Interface: Vlan-interface100      Circuit Id: 0000.0000.0004.01
State: Up      HoldTime: 30s      Type: L2              PRI: 64

```

```

# Display information about IS-IS interfaces of Switch A.
[SwitchA] display isis interface

```

Interface information for ISIS(1)

```

-----
Interface: Vlan-interface100

```

```

Id      IPV4.State    IPV6.State    MTU    Type    DIS
001     Up           Down         1497   L1/L2   No/No

```

Display information about IS-IS interfaces of Switch C.

```
[SwitchC] display isis interface
```

```

Interface information for ISIS(1)
-----

```

```
Interface: Vlan-interface100
```

```

Id      IPV4.State    IPV6.State    MTU    Type    DIS
001     Up           Down         1497   L1/L2   Yes/No

```

Display information about IS-IS interfaces of Switch D.

```
[SwitchD] display isis interface
```

```

Interface information for ISIS(1)
-----

```

```
Interface: Vlan-interface100
```

```

Id      IPV4.State    IPV6.State    MTU    Type    DIS
001     Up           Down         1497   L1/L2   No/Yes

```

By using the default DIS priority, Switch C is the Level-1 DIS, and Switch D is the Level-2 DIS. The pseudonodes of Level-1 and Level-2 are 0000.0000.0003.01 and 0000.0000.0004.01.

3. Configure the DIS priority of Switch A:

```
[SwitchA] interface vlan-interface 100
```

```
[SwitchA-Vlan-interface100] isis dis-priority 100
```

```
[SwitchA-Vlan-interface100] quit
```

Display IS-IS neighbors of Switch A.

```
[SwitchA] display isis peer
```

```

Peer information for ISIS(1)
-----

```

```

System Id: 0000.0000.0002
Interface: Vlan-interface100      Circuit Id: 0000.0000.0001.01
State: Up      HoldTime: 21s      Type: L1(L1L2)      PRI: 64

```

```

System Id: 0000.0000.0003
Interface: Vlan-interface100      Circuit Id: 0000.0000.0001.01
State: Up      HoldTime: 27s      Type: L1              PRI: 64

```

```

System Id: 0000.0000.0002
Interface: Vlan-interface100      Circuit Id: 0000.0000.0001.01
State: Up      HoldTime: 28s      Type: L2(L1L2)      PRI: 64

```

```

System Id: 0000.0000.0004
Interface: Vlan-interface100      Circuit Id: 0000.0000.0001.01
State: Up      HoldTime: 30s      Type: L2              PRI: 64

```

Display information about IS-IS interfaces of Switch A.

```
[SwitchA] display isis interface
```

```
Interface information for ISIS(1)
```

```
-----
```

```
Interface: Vlan-interface100
```

Id	IPV4.State	IPV6.State	MTU	Type	DIS
001	Up	Down	1497	L1/L2	Yes/Yes

After the DIS priority configuration, Switch A becomes the Level-1-2 DIS, and the pseudonode is 0000.0000.0001.01.

```
# Display information about IS-IS neighbors and interfaces of Switch C.
```

```
[SwitchC] display isis peer
```

```
Peer information for ISIS(1)
```

```
-----
```

```
System Id: 0000.0000.0002
```

Interface: Vlan-interface100	Circuit Id: 0000.0000.0001.01		
State: Up	HoldTime: 25s	Type: L1	PRI: 64

```
System Id: 0000.0000.0001
```

Interface: Vlan-interface100	Circuit Id: 0000.0000.0001.01		
State: Up	HoldTime: 7s	Type: L1	PRI: 100

```
[SwitchC] display isis interface
```

```
Interface information for ISIS(1)
```

```
-----
```

```
Interface: Vlan-interface100
```

Id	IPV4.State	IPV6.State	MTU	Type	DIS
001	Up	Down	1497	L1/L2	No/No

```
# Display information about IS-IS neighbors and interfaces of Switch D.
```

```
[SwitchD] display isis peer
```

```
Peer information for ISIS(1)
```

```
-----
```

```
System Id: 0000.0000.0001
```

Interface: Vlan-interface100	Circuit Id: 0000.0000.0001.01		
State: Up	HoldTime: 9s	Type: L2	PRI: 100

```
System Id: 0000.0000.0002
```

Interface: Vlan-interface100	Circuit Id: 0000.0000.0001.01		
State: Up	HoldTime: 28s	Type: L2	PRI: 64

```
[SwitchD] display isis interface
```

```
Interface information for ISIS(1)
```

```
-----
```

```
Interface: Vlan-interface100
```

Id	IPV4.State	IPV6.State	MTU	Type	DIS
----	------------	------------	-----	------	-----

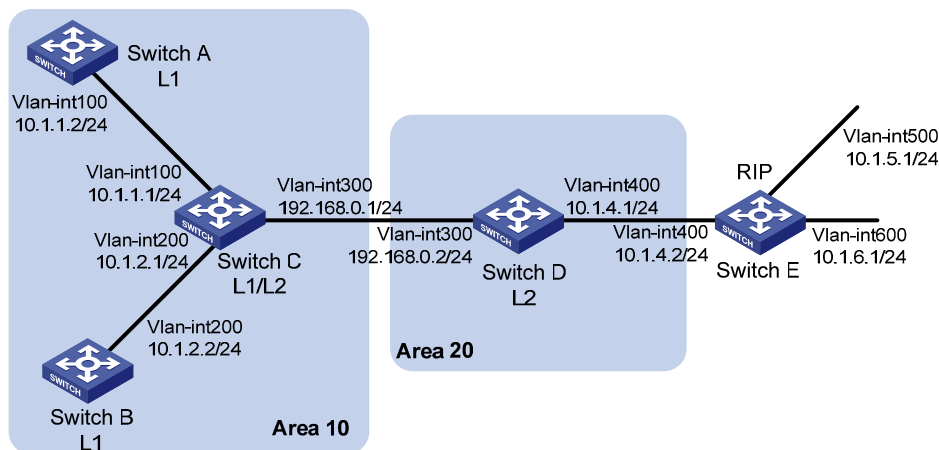
Configuring IS-IS route redistribution

Network requirements

As shown in [Figure 65](#), Switch A, Switch B, Switch C, and Switch D reside in the same AS. They use IS-IS to interconnect. Switch A and Switch B are Level-1 routers, Switch D is a Level-2 router, and Switch C is a Level-1-2 router.

Redistribute RIP routes into IS-IS on Switch D.

Figure 65 Network diagram for IS-IS route redistribution



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure IS-IS basic functions:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] isis 1
[SwitchA-isis-1] is-level level-1
[SwitchA-isis-1] network-entity 10.0000.0000.0001.00
[SwitchA-isis-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] isis enable 1
[SwitchA-Vlan-interface100] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] isis 1
[SwitchB-isis-1] is-level level-1
[SwitchB-isis-1] network-entity 10.0000.0000.0002.00
[SwitchB-isis-1] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] isis enable 1
[SwitchB-Vlan-interface200] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] isis 1
[SwitchC-isis-1] network-entity 10.0000.0000.0003.00
[SwitchC-isis-1] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] isis enable 1
[SwitchC-Vlan-interface200] quit
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] isis enable 1
[SwitchC-Vlan-interface100] quit
[SwitchC] interface vlan-interface 300
[SwitchC-Vlan-interface300] isis enable 1
[SwitchC-Vlan-interface300] quit
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] isis 1
[SwitchD-isis-1] is-level level-2
[SwitchD-isis-1] network-entity 20.0000.0000.0004.00
[SwitchD-isis-1] quit
[SwitchD] interface interface vlan-interface 300
[SwitchD-Vlan-interface300] isis enable 1
[SwitchD-Vlan-interface300] quit
```

Display IS-IS routing information on each switch.

```
[SwitchA] display isis route
```

Route information for ISIS(1)

ISIS(1) IPv4 Level-1 Forwarding Table

IPv4 Destination	IntCost	ExtCost	ExitInterface	NextHop	Flags
10.1.1.0/24	10	NULL	VLAN100	Direct	D/L/-
10.1.2.0/24	20	NULL	VLAN100	10.1.1.1	R/-/-
192.168.0.0/24	20	NULL	VLAN100	10.1.1.1	R/-/-
0.0.0.0/0	10	NULL	VLAN100	10.1.1.1	R/-/-

Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set

```
[SwitchC] display isis route
```

Route information for ISIS(1)

ISIS(1) IPv4 Level-1 Forwarding Table

IPv4 Destination	IntCost	ExtCost	ExitInterface	NextHop	Flags
10.1.1.0/24	10	NULL	VLAN100	Direct	D/L/-
10.1.2.0/24	10	NULL	VLAN200	Direct	D/L/-
192.168.0.0/24	10	NULL	VLAN300	Direct	D/L/-

Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set

ISIS(1) IPv4 Level-2 Forwarding Table

IPv4 Destination	IntCost	ExtCost	ExitInterface	NextHop	Flags
10.1.1.0/24	10	NULL	VLAN100	Direct	D/L/-
10.1.2.0/24	10	NULL	VLAN200	Direct	D/L/-
192.168.0.0/24	10	NULL	VLAN300	Direct	D/L/-

Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set

[SwitchD] display isis route

Route information for ISIS(1)

ISIS(1) IPv4 Level-2 Forwarding Table

IPv4 Destination	IntCost	ExtCost	ExitInterface	NextHop	Flags
192.168.0.0/24	10	NULL	VLAN300	Direct	D/L/-
10.1.1.0/24	20	NULL	VLAN300	192.168.0.1	R/-/-
10.1.2.0/24	20	NULL	VLAN300	192.168.0.1	R/-/-

Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set

3. Configure RIPv2 on Switch D and Switch E, and configure route redistribution from RIP to IS-IS on Switch D:

Configure RIPv2 on Switch D.

```
[SwitchD] rip 1
[SwitchD-rip-1] network 10.0.0.0
[SwitchD-rip-1] version 2
[SwitchD-rip-1] undo summary
```

Configure RIPv2 on Switch E.

```
[SwitchE] rip 1
[SwitchE-rip-1] network 10.0.0.0
[SwitchE-rip-1] version 2
[SwitchE-rip-1] undo summary
```

```
# Configure route redistribution from RIP to IS-IS on Switch D.
```

```
[SwitchD-rip-1] quit
[SwitchD] isis 1
[SwitchD-isis-1] import-route rip level-2
```

```
# Display IS-IS routing information on Switch C.
```

```
[SwitchC] display isis route
```

```
Route information for ISIS(1)
```

```
-----
ISIS(1) IPv4 Level-1 Forwarding Table
-----
```

IPv4 Destination	IntCost	ExtCost	ExitInterface	NextHop	Flags
10.1.1.0/24	10	NULL	VLAN100	Direct	D/L/-
10.1.2.0/24	10	NULL	VLAN200	Direct	D/L/-
192.168.0.0/24	10	NULL	VLAN300	Direct	D/L/-

```
Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set
```

```
-----
ISIS(1) IPv4 Level-2 Forwarding Table
-----
```

IPv4 Destination	IntCost	ExtCost	ExitInterface	NextHop	Flags
10.1.1.0/24	10	NULL	VLAN100	Direct	D/L/-
10.1.2.0/24	10	NULL	VLAN200	Direct	D/L/-
192.168.0.0/24	10	NULL	VLAN300	Direct	D/L/-
10.1.4.0/24	10	NULL	VLAN300	192.168.0.2	R/L/-
10.1.5.0/24	20	NULL	VLAN300	192.168.0.2	R/L/-
10.1.6.0/24	20	NULL	VLAN300	192.168.0.2	R/L/-

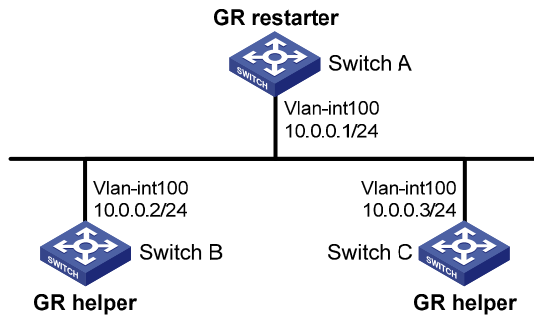
```
Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set
```

IS-IS Graceful Restart configuration example

Network requirements

Switch A, Switch B, and Switch C belong to the same IS-IS routing domain, as illustrated in [Figure 66](#).

Figure 66 Network diagram for IS-IS GR configuration



Configuration procedure

1. Configure IP addresses of the interfaces on each switch and configure IS-IS:
Follow Figure 66 to configure the IP address and subnet mask of each interface. (Details not shown.)
Configure IS-IS on the switches, ensuring that Switch A, Switch B and Switch C can communicate with each other at layer 3 and dynamic route update can be implemented among them with IS-IS. (Details not shown.)

2. Configure IS-IS Graceful Restart:

Enable IS-IS Graceful Restart on Switch A and configure the Graceful Restart Interval.

```
<SwitchA> system-view
[SwitchA] isis 1
[SwitchA-isis-1] graceful-restart
[SwitchA-isis-1] graceful-restart interval 150
[SwitchA-isis-1] return
```

Configurations for Switch B and Switch C are similar; therefore, details are not shown.

3. Verify the configuration:

After Router A establishes adjacencies with Router B and Router C, they begin to exchange routing information. Restart IS-IS on Router A, which enters the restart state and sends connection requests to its neighbors through the Graceful Restart mechanism to synchronize the LSDB. Using the **display isis graceful-restart status** command can display the IS-IS GR status on Router A.

Restart the IS-IS process on Switch A.

```
<SwitchA> reset isis all 1
Warning : Reset ISIS process? [Y/N]:y
```

Check the Graceful Restart status of IS-IS on Switch A.

```
<SwitchA> display isis graceful-restart status
Restart information for IS-IS(1)
```

IS-IS(1) Level-1 Restart Status

Restart Interval: 150

SA Bit Supported

Total Number of Interfaces = 1

Restart Status: RESTARTING

Number of LSPs Awaited: 3

T3 Timer Status:

Remaining Time: 140


```

T2 Timer Status:
    Remaining Time: 59

IS-IS(1) Level-2 Restart Status
Restart Interval: 150
SA Bit Supported
    Total Number of Interfaces = 1
Restart Status: RESTARTING
Number of LSPs Awaited: 3
T3 Timer Status:
    Remaining Time: 140
T2 Timer Status:
    Remaining Time: 59

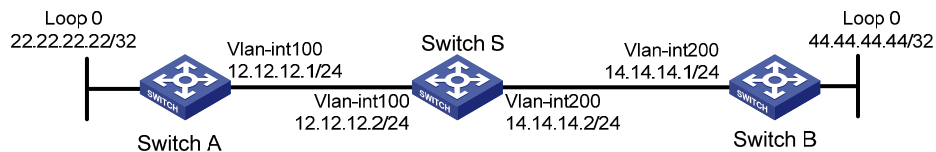
```

IS-IS NSR configuration example

Network requirements

Switch S, Switch A, and Switch B belong to the same IS-IS routing domain as illustrated in [Figure 67](#). Switch S is an IRF fabric, which is connected to Switch A and Switch B over aggregated links. Enable IS-IS NSR on Switch S to ensure correct routing when a master/slave switchover occurs on Switch S.

Figure 67 Network diagram for IS-IS NSR configuration



Configuration procedure

- Configure IP addresses for the interfaces on each switch and configure IS-IS:
Follow [Figure 67](#) to configure the IP address and subnet mask of each interface on the switches. (Details not shown.)
Configure IS-IS on the switches, ensuring that Switch S, Switch A, and Switch B can communicate with each other at Layer 3 and dynamic route update can be implemented among them with IS-IS. (Details not shown.)
- Configure IS-IS NSR:
Enable IS-IS NSR on Switch S.

```

<SwitchS> system-view
[SwitchS] isis 1
[SwitchS-isis-1] non-stop-routing
[SwitchS-isis-1] non-stop-routing interval 30
[SwitchS-isis-1] return

```
- Verify the configuration:
After Switch S establishes neighbor relationships with Switch A and Switch B, they start to exchange routing information. After network convergence, perform a master/slave switchover on Switch S. During the switchover period, use the **display isis peer** command to check the neighbor relationship between Switch A and Switch S and between Switch B and Switch S; use the **display**

isis route command to check if routes from Switch A to the loopback interface on Switch B and from Switch B to the loopback interface on Switch A exist.

When a master/slave switchover occurs on Switch S, display IS-IS neighbors and routes on Switch A.

```
<SwitchA> display isis peer
```

```
Peer information for ISIS(1)
```

```
-----  
System Id: 0000.0000.0001  
Interface: vlan100          Circuit Id: 0000.0000.0001.01  
State: Up    HoldTime: 25s  Type: L1(L1L2)    PRI: 64
```

```
System Id: 0000.0000.0001  
Interface: vlan100          Circuit Id: 0000.0000.0001.01  
State: Up    HoldTime: 27s  Type: L2(L1L2)    PRI: 64
```

```
<SwitchA> display isis route
```

```
Route information for ISIS(1)
```

```
-----  
ISIS(1) IPv4 Level-1 Forwarding Table
```

```
-----  
IPV4 Destination    IntCost    ExtCost    ExitInterface    NextHop    Flags  
-----  
12.12.12.0/24       10         NULL       vlan100          Direct     D/L/-  
22.22.22.22/32      10         NULL       Loop0            Direct     D/-/-  
14.14.14.0/32       10         NULL       vlan100          12.12.12.2  R/L/-  
44.44.44.44/32      10         NULL       vlan100          12.12.12.2  R/L/-
```

```
Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set
```

```
ISIS(1) IPv4 Level-2 Forwarding Table
```

```
-----  
IPV4 Destination    IntCost    ExtCost    ExitInterface    NextHop    Flags  
-----  
12.12.12.0/24       10         NULL       vlan100          Direct     D/L/-  
22.22.22.22/32      10         NULL       Loop0            Direct     D/-/-  
14.14.14.0/32       10         NULL  
44.44.44.44/32      10         NULL
```

```
Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set
```

Display IS-IS neighbors and routes on Switch B.

```
<SwitchB> display isis peer
```

```
Peer information for ISIS(1)
```

```

System Id: 0000.0000.0001
Interface: vlan200          Circuit Id: 0000.0000.0001.01
State: Up      HoldTime: 25s  Type: L1(L1L2)    PRI: 64

System Id: 0000.0000.0001
Interface: vlan200          Circuit Id: 0000.0000.0001.01
State: Up      HoldTime: 27s  Type: L2(L1L2)    PRI: 64
<SwitchB> display isis route
Route information for ISIS(1)
-----
ISIS(1) IPv4 Level-1 Forwarding Table
-----
IPV4 Destination      IntCost    ExtCost  ExitInterface  NextHop      Flags
-----
14.14.14.0/24         10         NULL     vlan200        Direct       D/L/-
44.44.44.44/32        10         NULL     Loop0          Direct       D/-/-
12.12.12.0/32         10         NULL     vlan200        14.14.14.4  R/L/-
22.22.22.22/32        10         NULL     vlan200        14.14.14.4  R/L/-

Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set

```

```

ISIS(1) IPv4 Level-2 Forwarding Table
-----
IPV4 Destination      IntCost    ExtCost  ExitInterface  NextHop      Flags
-----
14.14.14.0/24         10         NULL     vlan200        Direct       D/L/-
44.44.44.44/32        10         NULL     Loop0          Direct       D/-/-
12.12.12.0/32         10         NULL
22.22.22.22/32        10         NULL

Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set

```

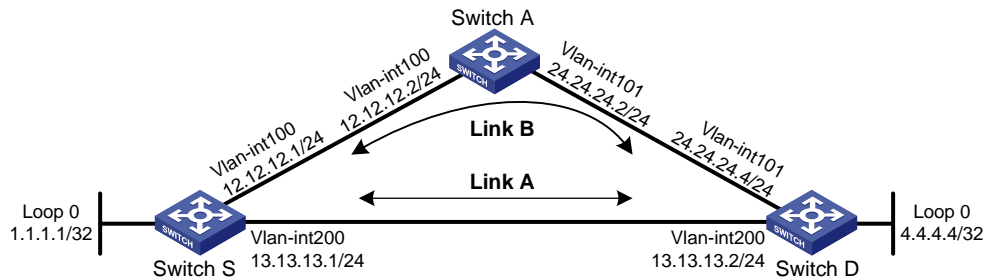
The output shows that the neighbor relationships and routing information on Switch A and Switch B have not changed. The neighbors cannot sense the switchover on Switch S.

IS-IS FRR configuration example

Network requirements

Switch S, Switch A, and Switch D belong to the same IS-IS routing domain as illustrated in [Figure 68](#). Configure IS-IS FRR so that when the link between Switch S and Switch D fails, traffic can be switched to Link B immediately.

Figure 68 Network diagram for IS-IS FRR configuration



Configuration procedure

1. Configure IP addresses for the interfaces on each switch and configure IS-IS:

Follow [Figure 68](#) to configure the IP address and subnet mask of each interface on the switches. (Details not shown.)

Configure IS-IS on the switches, ensuring that Switch A, Switch D, and Switch S can communicate with each other at Layer 3. (Details not shown.)

2. Configure IS-IS FRR:

You can enable IS-IS FRR to either automatically calculate a backup next hop, or to designate a backup next hop by using a referenced routing policy.

Method I: Enable IS-IS FRR to automatically calculate a backup next hop.

Configure Switch S.

```
<SwitchS> system-view
[SwitchS] bfd echo-source-ip 1.1.1.1
[SwitchS] isis 1
[SwitchS-isis-1] fast-reroute auto
[SwitchS-isis-1] quit
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] bfd echo-source-ip 4.4.4.4
[SwitchD] isis 1
[SwitchD-isis-1] fast-reroute auto
[SwitchD-isis-1] quit
```

Method II: Enable IS-IS FRR to designate a backup next hop by using a referenced routing policy.

Configure Switch S.

```
<SwitchS> system-view
[SwitchS] bfd echo-source-ip 1.1.1.1
[SwitchS] ip ip-prefix abc index 10 permit 4.4.4.4 32
[SwitchS] route-policy frr permit node 10
[SwitchS-route-policy] if-match ip-prefix abc
[SwitchS-route-policy] apply fast-reroute backup-interface vlan-interface 100
backup-nexthop 12.12.12.2
[SwitchS-route-policy] quit
[SwitchS] isis 1
[SwitchS-isis-1] fast-reroute route-policy frr
[SwitchS-isis-1] quit
```

Configure Switch D.

```

<SwitchD> system-view
[SwitchD] bfd echo-source-ip 4.4.4.4
[SwitchD] ip ip-prefix abc index 10 permit 1.1.1.1 32
[SwitchD] route-policy frr permit node 10
[SwitchD-route-policy] if-match ip-prefix abc
[SwitchD-route-policy] apply fast-reroute backup-interface vlan-interface 101
backup-nexthop 24.24.24.2
[SwitchD-route-policy] quit
[SwitchD] isis 1
[SwitchD-isis-1] fast-reroute route-policy frr
[SwitchD-isis-1] quit

```

3. Verify the configuration:

Display route 4.4.4.4/32 on Switch S and you can view the backup next hop information.

```

[SwitchS] display ip routing-table 4.4.4.4 verbose
Routing Table : Public
Summary Count : 1

```

```

Destination: 4.4.4.4/32
  Protocol: ISIS                Process ID: 1
  Preference: 10                Cost: 10
  IpPrecedence:                QoSLeId:
  NextHop: 13.13.13.2          Interface: Vlan-interface200
  BkNextHop: 12.12.12.2       BkInterface: Vlan-interface100
  RelyNextHop: 0.0.0.0        Neighbor : 0.0.0.0
  Tunnel ID: 0x0              Label: NULL
  BKTunnel ID: 0x0           BKLabel: NULL
  State: Active Adv          Age: 00h01m27s
  Tag: 0

```

Display route 1.1.1.1/32 on Switch D. You can find the backup next hop information.

```

[SwitchD] display ip routing-table 1.1.1.1 verbose
Routing Table : Public
Summary Count : 1

```

```

Destination: 1.1.1.1/32
  Protocol: ISIS                Process ID: 1
  Preference: 10                Cost: 10
  IpPrecedence:                QoSLeId:
  NextHop: 13.13.13.1          Interface: Vlan-interface200
  BkNextHop: 24.24.24.2       BkInterface: Vlan-interface101
  RelyNextHop: 0.0.0.0        Neighbor : 0.0.0.0
  Tunnel ID: 0x0              Label: NULL
  BKTunnel ID: 0x0           BKLabel: NULL
  State: Active Adv          Age: 00h01m27s
  Tag: 0

```

IS-IS authentication configuration example

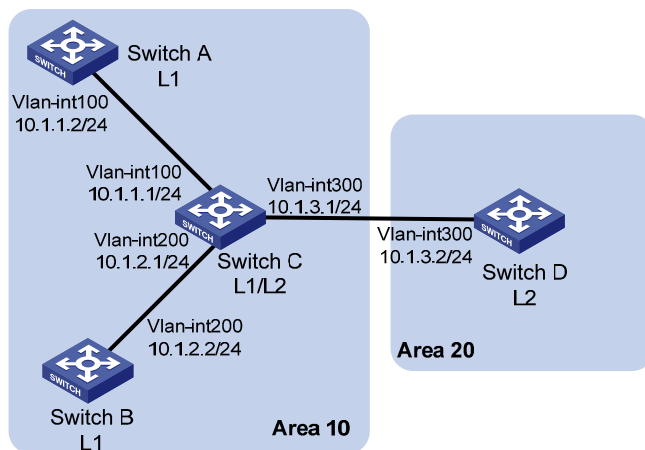
Network requirements

As shown in Figure 69, Switch A, Switch B, Switch C and Switch D reside in the same IS-IS routing domain.

Switch A, Switch B, and Switch C belong to Area 10, and Switch D belongs to Area 20.

Configure relationship authentication between neighbors. Configure area authentication in Area 10 to prevent untrusted routes from entering into the area. Configure routing domain authentication on Switch C and Switch D to prevent untrusted routes from entering the routing domain.

Figure 69 Network diagram for IS-IS authentication configuration



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure IS-IS basic functions:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] isis 1
[SwitchA-isis-1] network-entity 10.0000.0000.0001.00
[SwitchA-isis-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] isis enable 1
[SwitchA-Vlan-interface100] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] isis 1
[SwitchB-isis-1] network-entity 10.0000.0000.0002.00
[SwitchB-isis-1] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] isis enable 1
[RouterB--Vlan-interface200] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] isis 1
```

```
[SwitchC-isis-1] network-entity 10.0000.0000.0003.00
[SwitchC-isis-1] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] isis enable 1
[SwitchC-Vlan-interface200] quit
[SwitchC] interface vlan-interface 300
[SwitchC-Vlan-interface300] isis enable 1
[SwitchC-Vlan-interface300] quit
[SwitchC] interface vlan-interface 300
[SwitchC-Vlan-interface300] isis enable 1
[SwitchC-Vlan-interface300] quit
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] isis 1
[SwitchD-isis-1] network-entity 20.0000.0000.0001.00
[SwitchD-isis-1] quit
[SwitchD] interface vlan-interface 300
[SwitchD-Vlan-interface300] isis enable 1
[SwitchD-Vlan-interface300] quit
```

3. Configure neighbor relationship authentication between neighbors:

Specify the MD5 authentication mode and password **eRq** on VLAN-interface 100 of Switch A and on VLAN-interface 100 of Switch C.

```
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] isis authentication-mode md5 eRg
[SwitchA-Vlan-interface100] quit
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] isis authentication-mode md5 eRg
[SwitchC-Vlan-interface100] quit
```

Specify the MD5 authentication mode and password **t5Hr** on VLAN-interface 200 of Switch B and on VLAN-interface 200 of Switch C.

```
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] isis authentication-mode md5 t5Hr
[SwitchB-Vlan-interface200] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] isis authentication-mode md5 t5Hr
[SwitchC-Vlan-interface200] quit
```

Specify the MD5 authentication mode and password **hSec** on VLAN-interface 300 of Switch D and on VLAN-interface 300 of Switch C.

```
[SwitchC] interface vlan-interface 300
[SwitchC-Vlan-interface300] isis authentication-mode md5 hSec
[SwitchC-Vlan-interface300] quit
[SwitchD] interface vlan-interface 300
[SwitchD-Vlan-interface300] isis authentication-mode md5 hSec
[SwitchD-Vlan-interface300] quit
```

4. Configure area authentication. Specify the MD5 authentication mode and password **10Sec** on Switch A, Switch B, and Switch C:

```
[SwitchA] isis 1
```

```

[SwitchA-isis-1] area-authentication-mode md5 10Sec
[SwitchA-isis-1] quit
[SwitchB] isis 1
[SwitchB-isis-1] area-authentication-mode md5 10Sec
[SwitchB-isis-1] quit
[SwitchC] isis 1
[SwitchC-isis-1] area-authentication-mode md5 10Sec
[SwitchC-isis-1] quit

```

5. Configure routing domain authentication. Specify the MD5 authentication mode and password **1020Sec** on Switch C and Switch D:

```

[SwitchC] isis 1
[SwitchC-isis-1] domain-authentication-mode md5 1020Sec
[SwitchC-isis-1] quit
[SwitchD] isis 1
[SwitchD-isis-1] domain-authentication-mode md5 1020Sec

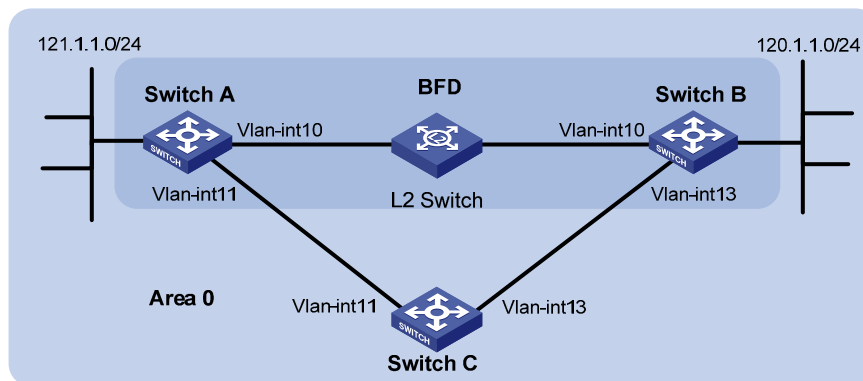
```

Configuring BFD for IS-IS

Network requirements

- As shown in [Figure 70](#), IS-IS is enabled on Switch A, Switch B and Switch C that are reachable to each other at the network layer.
- After the link over which Switch A and Switch B communicate through the Layer-2 switch fails, BFD can quickly detect the failure and notify IS-IS of the failure. Switch A and Switch B then communicate through Switch C.

Figure 70 Network diagram for BFD configuration on an IS-IS link



Device	Interface	IP address	Device	Interface	IP address
Switch A	Vlan-int10	10.1.0.102/24	Switch B	Vlan-int10	10.1.0.100/24
	Vlan-int11	11.1.1.1/24		Vlan-int13	13.1.1.1/24
Switch C	Vlan-int11	11.1.1.2/24			
	Vlan-int13	13.1.1.2/24			

Configuration procedure

1. Configure IP addresses for interfaces (Details not shown.).
2. Configure IS-IS basic functions.
Configure Switch A.


```

<SwitchA> system-view
[SwitchA] isis
[SwitchA-isis-1] network-entity 10.0000.0000.0001.00
[SwitchA-isis-1] quit
[SwitchA] interface vlan-interface 10
[SwitchA-Vlan-interface10] isis enable
[SwitchA-Vlan-interface10] quit
[SwitchA] interface vlan-interface 11
[SwitchA-Vlan-interface11] isis enable
[SwitchA-Vlan-interface11] quit

```

Configure Switch B.

```

<SwitchB> system-view
[SwitchB] isis
[SwitchB-isis-1] network-entity 10.0000.0000.0002.00
[SwitchB-isis-1] quit
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] isis enable
[SwitchB-Vlan-interface10] quit
[SwitchB] interface vlan-interface 13
[SwitchB-Vlan-interface13] isis enable
[SwitchB-Vlan-interface13] quit

```

Configure Switch C.

```

<SwitchC> system-view
[SwitchC] isis
[SwitchC-isis-1] network-entity 10.0000.0000.0003.00
[SwitchC-isis-1] quit
[SwitchC] interface vlan-interface 11
[SwitchC-Vlan-interface11] isis enable
[SwitchC-Vlan-interface11] quit
[SwitchC] interface vlan-interface 13
[SwitchC-Vlan-interface13] isis enable
[SwitchC-Vlan-interface13] quit

```

3. Configure BFD parameters:

Enable BFD on Switch A and configure BFD parameters.

```

[SwitchA] bfd session init-mode active
[SwitchA] interface vlan-interface 10
[SwitchA-Vlan-interface10] isis bfd enable
[SwitchA-Vlan-interface10] bfd min-receive-interval 500
[SwitchA-Vlan-interface10] bfd min-transmit-interval 500
[SwitchA-Vlan-interface10] bfd detect-multiplier 7

```

Enable BFD on Switch B and configure BFD parameters.

```

[SwitchB] bfd session init-mode active
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] isis bfd enable
[SwitchB-Vlan-interface10] bfd min-receive-interval 500
[SwitchB-Vlan-interface10] bfd min-transmit-interval 500
[SwitchB-Vlan-interface10] bfd detect-multiplier 8

```

```
[SwitchB-Vlan-interface10] return
```

4. Verify the configuration:

The following configurations are made on Switch A. Configurations for Switch B are similar. (Details not shown.)

Display the BFD information of Switch A.

```
<SwitchA> display bfd session
Total Session Num: 1          Init Mode: Active
Session Working Under Ctrl Mode:
LD/RD      SourceAddr      DestAddr      State Holdtime Interface
3/1        10.1.0.102      10.1.0.100   Up    1700ms  vlan10
```

Display route 120.1.1.0/24 on Switch A, and you can see that Switch A and Switch B communicate through the Layer-2 switch.

```
<SwitchA> display ip routing-table 120.1.1.0 verbose
Routing Table : Public
Summary Count : 2
  Destination: 120.1.1.0/24
    Protocol: ISIS          Process ID: 0
    Preference: 0          Cost: 2
    IpPrecedence:          QoS LcId:
      NextHop: 192.168.0.100  Interface: Vlan-interface10
      BkNextHop: 0.0.0.0      BkInterface:
      RelyNextHop: 0.0.0.0    Neighbor : 0.0.0.0
      Tunnel ID: 0x0          Label: NULL
      BKTunnel ID: 0x0        BKLabel: NULL
      State: Active Adv      Age: 00h58m10s
      Tag: 0
  Destination: 120.1.1.0/24
    Protocol: ISIS          Process ID: 1
    Preference: 10         Cost: 4
    IpPrecedence:          QoS LcId:
      NextHop: 10.1.1.100    Interface: Vlan-interface11
      BkNextHop: 0.0.0.0      BkInterface:
      RelyNextHop: 0.0.0.0    Neighbor : 0.0.0.0
      Tunnel ID: 0x0          Label: NULL
      BKTunnel ID: 0x0        BKLabel: NULL
      State: Invalid Adv     Age: 00h58m05s
      Tag: 0
```

Enable debugging on Switch A.

```
<SwitchA> debugging isis bfd-event
<SwitchA> terminal debugging
```

When the link between Switch B and the Layer-2 switch fails, BFD can quickly detect the failure.

```
#Aug 8 14:54:05:362 2008 SwitchA IFNET/4/INTERFACE UPDOWN:
Trap 1.3.6.1.6.3.1.1.5.3<linkDown>: Interface 983041 is Down, ifAdminStatus is 1,
ifOperStatus is 2
#Aug 8 14:54:05:363 2008 SwitchA
ISIS/4/ADJ_CHANGE:TrapID(1.3.6.1.2.1.138.0.17<isisAdjacencyChange>), ISIS Level-2
Adjacency IN Circuit-983041 State Change.
```

```

#Aug  8 14:54:05:364 2008 SwitchA
ISIS/4/ADJ_CHANGE:TrapID(1.3.6.1.2.1.138.0.17<isisAdjacencyChange>), ISIS Level-1
Adjacency IN Circuit-983041 State Change.
%Aug  8 14:54:05:365 2008 SwitchA IFNET/4/LINK UPDOWN: Vlan-interface10 link status
is DOWN
%Aug  8 14:54:05:366 2008 SwitchA IFNET/4/UPDOWN: Line protocol on the interface
Vlan-interface10 is DOWN
%Aug  8 14:54:05:367 2008 SwitchA ISIS/4/ADJLOG:ISIS-1-ADJCHANGE: Adjacency To
0000.0000.0002 (vlan10) DOWN, Level-2 Circuit Down.
%Aug  8 14:54:05:367 2008 SwitchA ISIS/4/ADJLOG:ISIS-1-ADJCHANGE: Adjacency To
0000.0000.0002 (vlan10) DOWN, Level-2 Adjacency clear.
%Aug  8 14:54:05:368 2008 SwitchA ISIS/4/ADJLOG:ISIS-1-ADJCHANGE: Adjacency To
0000.0000.0002 (vlan10) DOWN, Level-1 Circuit Down.
%Aug  8 14:54:05:369 2008 SwitchA ISIS/4/ADJLOG:ISIS-1-ADJCHANGE: Adjacency To
0000.0000.0002 (vlan10) DOWN, Level-1 Adjacency clear.
*Aug  8 14:54:05:370 2008 SwitchA ISIS/6/ISIS:
  ISIS-1-BFD: Success to send msg. Msg type 1 delete session. IfPhyIndex: 5 ,DstIPAddr:
10.1.0.100 , SrcIPAddr:10.1.0.102. NeighborType:Level-2.

*Aug  8 14:54:05:370 2008 SwitchA ISIS/6/ISIS:
  ISIS-1-BFD: Success to send msg. Msg type 1 delete session. IfPhyIndex: 5 ,DstIPAddr:
10.1.0.100 , SrcIPAddr:10.1.0.102. NeighborType:Level-1.
# Display the BFD information of Switch A. Switch A has removed its neighbor relationship with
Switch B and no information is output.
<SwitchA> display bfd session

# Display route 120.1.1.0/24 on Switch A, and you can see that Switch A and Switch B
communicate through Switch C.
<SwitchA> display ip routing-table 120.1.1.0 verbose
Routing Table : Public
Summary Count : 2
  Destination: 120.1.1.0/24
    Protocol: ISIS                Process ID: 1
    Preference: 10                Cost: 4
    IpPrecedence:                QoSLeId:
  NextHop: 10.1.1.100            Interface: Vlan-interface11
  BkNextHop: 0.0.0.0            BkInterface:
  RelyNextHop: 0.0.0.0          Neighbor : 0.0.0.0
  Tunnel ID: 0x0                Label: NULL
  BKTunnel ID: 0x0              BKLabel: NULL
  State: Active Adv              Age: 00h58m10s
  Tag: 0

```

Configuring BGP

Hardware compatibility

The HP 5500 SI Switch Series does not support BGP.

BGP overview

The Border Gateway Protocol (BGP) is a dynamic inter-AS Exterior Gateway Protocol. The three early BGP versions are BGP-1 (RFC 1105), BGP-2 (RFC 1163) and BGP-3 (RFC 1267). The current version is BGP-4 (RFC 4271), and is the Internet exterior gateway protocol. The term "router" in this chapter refers to both routers and Layer 3 switches, and BGP refers to BGP-4 in this chapter.

BGP has the following characteristics:

- Focuses on the control of route propagation and the selection of optimal routes rather than the route discovery and calculation, which makes BGP, an exterior gateway protocol different from interior gateway protocols such as OSPF and RIP.
- Uses TCP to enhance reliability.
- Supports CIDR.
- Reduces bandwidth consumption by advertising only incremental updates and is applicable to advertising a great amount of routing information on the Internet.
- Eliminates routing loops completely by adding AS path information to BGP route advertisements.
- Provides abundant policies to implement flexible route filtering and selection.
- Provides good scalability.

A router advertising BGP messages is called a "BGP speaker". It establishes peer relationships with other BGP speakers to exchange routing information. When a BGP speaker receives a new route or a route better than the current one from another AS, it will advertise the route to all the other BGP peers in the local AS.

To simplify configuration, multiple peers using an identical policy can be organized as a peer group.

BGP runs on a router in either of the following modes:

- IBGP (internal BGP)
- EBGP (external BGP)

BGP is called "IBGP" when it runs within an AS, and is called "EBGP" when it runs between ASs.

BGP messages formats

Header

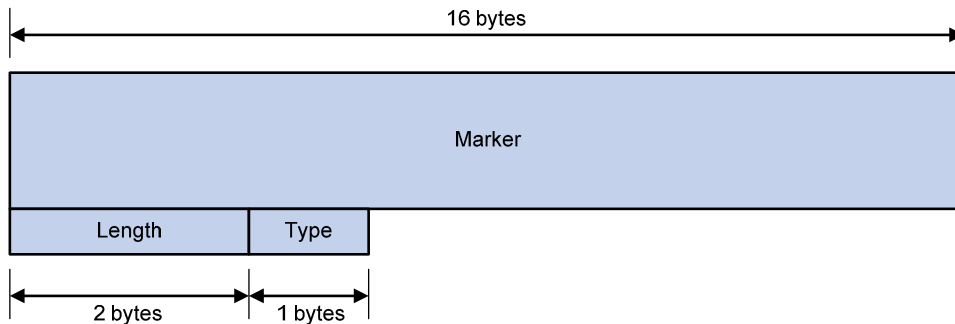
BGP has the following types of messages:

- Open
- Update
- Notification

- Keepalive
- Route-refresh

They have the same header.

Figure 71 BGP message header

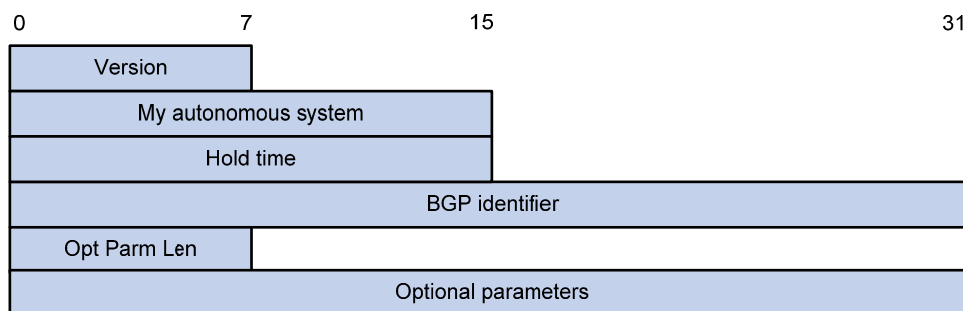


- **Marker**—The 16-byte field is used to delimit BGP messages. The Marker must be all ones.
- **Length**—The two-byte unsigned integer indicates the total length of the message.
- **Type**—This one-byte unsigned integer indicates the type code of the message. The following type codes are defined: 1–Open, 2–Update, 3–Notification, 4–Keepalive, and 5–Route-refresh. The former four listed codes are defined in RFC 1771, and the last listed code is defined in RFC 2918.

Open

After a TCP connection is established, the first message sent by each side is an open message for peer relationship establishment.

Figure 72 BGP open message format



Major fields of the BGP open message are as follows:

- **Version**—This one-byte unsigned integer indicates the protocol version number. The current BGP version is 4.
- **My autonomous system**—This two-byte unsigned integer indicates the autonomous system number of the sender.
- **Hold time**—When establishing a peer relationship, two parties negotiate an identical hold time. If no Keepalive or Update is received from a peer within the hold time, the BGP connection is considered down.
- **BGP identifier**—An IP address that identifies the BGP router.
- **Opt Parm Len (Optional Parameters Length)**—Length of optional parameters, which is set to 0 if no optional parameter is available.

- **Optional parameters**—Used for multiprotocol extensions and other functions.

Update

The update messages are used to exchange routing information between peers. It can advertise feasible routes or remove multiple unfeasible routes.

Figure 73 BGP update message format

Unfeasible routes length	2 Octets
Withdrawn routes	N Octets
Total path attribute length	2 Octets
Path attributes	N Octets
NLRI	N Octets

Each update message can advertise a group of feasible routes with identical attributes, and the routes are contained in the network layer reachability information (NLRI) field. The path attributes field carries the attributes of these routes. Each update message can also carry multiple withdrawn routes in the withdrawn routes field.

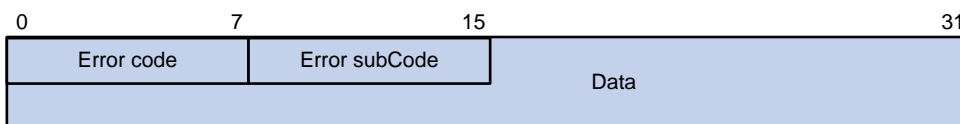
Major fields of the BGP update message are as follows:

- **Unfeasible routes length**—The total length of the withdrawn routes field in bytes. A value of 0 indicates no route is withdrawn from service, nor is the withdrawn routes field present in this update message.
- **Withdrawn routes**—This is a variable length field that contains a list of withdrawn IP prefixes.
- **Total path attribute length**—Total length of the path attributes field in bytes. A value of 0 indicates that no NLRI field is present in this update message.
- **Path attributes**—List of path attributes related to NLRI. Each path attribute is a triple <attribute type, attribute length, attribute value> of variable length. BGP uses these attributes to avoid routing loops, and perform routing and protocol extensions.
- **NLRI**— Each feasible route is represented as <length, prefix>.

Notification

A notification message is sent when an error is detected. The BGP connection is closed immediately after sending it.

Figure 74 BGP notification message format



Major fields of the BGP notification message are as follows:

- **Error code**—Type of notification.
- **Error subcode**—Specific information about the nature of the reported error.
- **Data**—Used to diagnose the reason for the notification. The contents of the data field depend on the error code and error subcode. Erroneous data can be recorded in the data field. The data field length is variable.

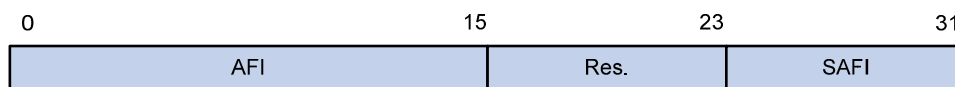
Keepalive

Keepalive messages are sent between peers to maintain connectivity. Its format contains only the message header.

Route-refresh

A route-refresh message is sent to a peer to request the specified address family routing information.

Figure 75 BGP route-refresh message format



- **AFI**—Address family identifier.
- **Res**—Reserved; set to 0.
- **SAFI**—Subsequent address family identifier.

BGP path attributes

Path attributes classification

BGP path attributes are a group of parameters encapsulated in the path attributes field of update messages. They give detailed route attributes information that can be used for route filtering and selection.

Path attributes fall into the following categories:

- **Well-known mandatory**—Must be recognized by all BGP routers and be included in every Update message. Routing information errors occur without this attribute.
- **Well-known discretionary**—Can be recognized by all BGP routers and optionally included in every Update message as needed.
- **Optional transitive**—Transitive attribute between ASs. A BGP router not supporting this attribute can still receive routes with this attribute and advertise them to other peers.
- **Optional non-transitive**—If a BGP router does not support this attribute, it will not advertise routes with this attribute.

The usage of each BGP path attribute is described in the following table.

Table 7 Usage of BGP path attributes

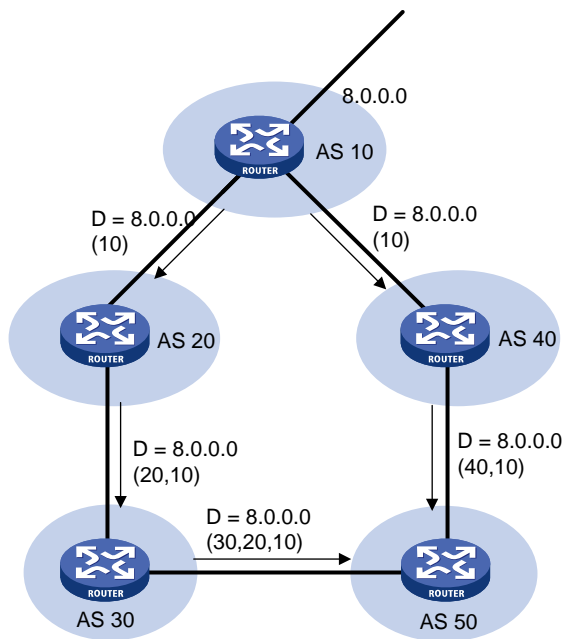
Name	Category
ORIGIN	Well-known mandatory
AS_PATH	Well-known mandatory
NEXT_HOP	Well-known mandatory
LOCAL_PREF	Well-known discretionary
ATOMIC_AGGREGATE	Well-known discretionary
AGGREGATOR	Optional transitive
COMMUNITY	Optional transitive
MULTI_EXIT_DISC (MED)	Optional non-transitive
ORIGINATOR_ID	Optional non-transitive

Name	Category
CLUSTER_LIST	Optional non-transitive

Usage of BGP path attributes

- **ORIGIN**
 ORIGIN is a well-known mandatory attribute that defines the origin of routing information (how a route became a BGP route). This attribute has the following types:
 - **IGP**—Has the highest priority. Routes added to the BGP routing table using the **network** command have the IGP attribute.
 - **EGP**—Has the second highest priority. Routes obtained via EGP have the EGP attribute.
 - **INCOMPLETE**—Has the lowest priority. The source of routes with this attribute is unknown, which does not mean such routes are unreachable. The routes redistributed from other routing protocols have the INCOMPLETE attribute.
- **AS_PATH**
 AS_PATH is a well-known mandatory attribute. This attribute identifies the autonomous systems through which routing information carried in this Update message has passed. When a route is advertised from the local AS to another AS, each passed AS number is added into the AS_PATH attribute, so the receiver can determine ASs to route the message back. The number of the AS closest to the receiver's AS is leftmost, as shown in [Figure 76](#):

Figure 76 AS_PATH attribute



Generally, a BGP router does not receive routes containing the local AS number to avoid routing loops.

NOTE:

The current implementation supports using the **peer allow-as-loop** command to receive routes containing the local AS number in order to meet special requirements.

Use the AS_PATH attribute for route selection and filtering. BGP gives priority to the route with the shortest AS_PATH length, if other factors are the same. As shown in [Figure 76](#), the BGP router in AS50 gives priority to the route passing AS40 for sending data to the destination 8.0.0.0.

In some applications, you can apply a routing policy to control BGP route selection by modifying the AS_PATH length.

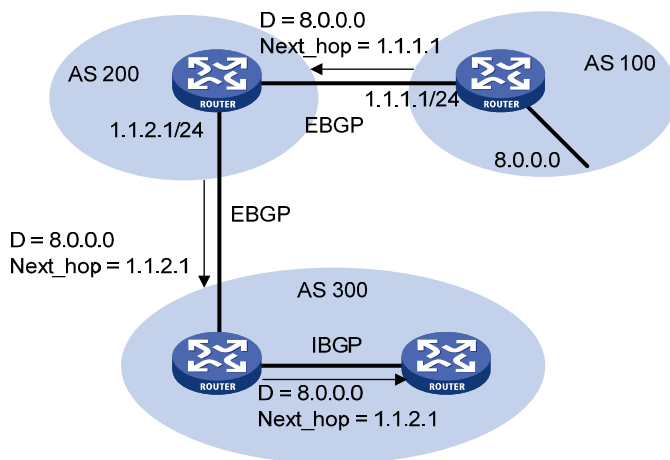
By configuring an AS path filtering list, you can filter routes based on AS numbers contained in the AS_PATH attribute.

- NEXT_HOP

Different from IGP, the NEXT_HOP attribute may not be the IP address of a directly connected router. It involves the following types of values, as shown in [Figure 77](#).

- When advertising a self-originated route to an EBGP peer, a BGP speaker sets the NEXT_HOP for the route to the address of its sending interface.
- When sending a received route to an EBGP peer, a BGP speaker sets the NEXT_HOP for the route to the address of the sending interface.
- When sending a route received from an EBGP peer to an IBGP peer, a BGP speaker does not modify the NEXT_HOP attribute. If load-balancing is configured, the NEXT_HOP attribute of the equal-cost routes is modified. For load-balancing information, see "[BGP route selection.](#)"

Figure 77 NEXT_HOP attribute

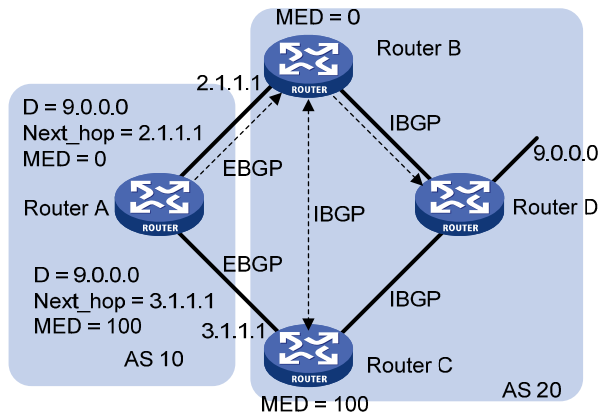


- MED (MULTI_EXIT_DISC)

The MED attribute is exchanged between two neighboring ASs, each of which does not advertise the attribute to any other AS.

Similar to metrics used by IGP, MED is used to determine the best route for traffic going into an AS. When a BGP router obtains multiple routes to the same destination, but with different next hops, it considers the route with the smallest MED value the best route given that other conditions are the same. As shown in [Figure 78](#), traffic from AS10 to AS20 travels through Router B that is selected according to MED.

Figure 78 MED attribute



In general, BGP compares MEDs of routes received from the same AS only.

NOTE:

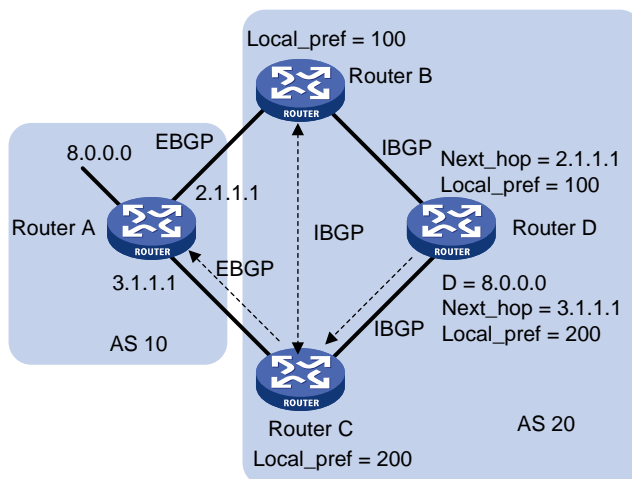
The current implementation supports using the **compare-different-as-med** command to force BGP to compare MED values of routes received from different ASs.

- LOCAL_PREF

The LOCAL_PREF attribute is exchanged between IBGP peers only; therefore, it is not advertised to any other AS. It indicates the priority of a BGP router.

LOCAL_PREF is used to determine the best route for traffic leaving the local AS. When a BGP router obtains from several IBGP peers multiple routes to the same destination, but with different next hops, it considers the route with the highest LOCAL_PREF value as the best route. As shown in Figure 79, traffic from AS20 to AS10 travels through Router C that is selected according to LOCAL_PREF.

Figure 79 LOCAL_PREF attribute



- COMMUNITY

The COMMUNITY attribute is a group of specific data. A route can carry one or more COMMUNITY attribute values (each of which is represented by a four-byte integer). The receiving router processes the route (for example, determining whether to advertise the route and the scope for advertising the route) based on the COMMUNITY attribute values. This simplifies routing policy

usage and facilitates management and maintenance. Well-known community attributes are as follows:

- **Internet**—By default, all routes belong to the Internet community. Routes with this attribute can be advertised to all BGP peers.
- **No_Export**—After received, routes with this attribute cannot be advertised out the local AS or out the local confederation, but can be advertised to other sub-ASs in the confederation. For confederation information, see "[Settlements for problems in large scale BGP networks.](#)"
- **No_Advertise**—After received, routes with this attribute cannot be advertised to other BGP peers.
- **No_Export_Subconfed**—After received, routes with this attribute cannot be advertised out the local AS or other ASs in the local confederation.

BGP route selection

Route selection rules

BGP discards routes with unreachable NEXT_HOPs. If multiple routes to the same destination are available, BGP selects the best route in the following sequence:

1. The route with the highest Preferred_value
2. The route with the highest LOCAL_PREF
3. The route originated by the local router
4. The route with the shortest AS-PATH
5. The IGP, EGP, or INCOMPLETE route in turn
6. The route with the lowest MED value
7. The route learned from EBGP, confederation, or IBGP in turn
8. The route with the smallest next hop metric
9. The route with the shortest CLUSTER_LIST
10. The route with the smallest ORIGINATOR_ID
11. The route advertised by the router with the smallest router ID
12. The route advertised by the peer with the lowest IP address

CLUSTER_IDs of route reflectors form a CLUSTER_LIST. If a route reflector receives a route that contains its own CLUSTER ID in the CLUSTER_LIST, the router discards the route to avoid routing loops.

If load balancing is configured, the system selects available routes to implement load balancing.

Route selection with BGP load balancing

The next hop of a BGP route may not be directly connected. One of the reasons is next hops in routing information exchanged between IBGPs are not modified. The BGP router needs to find the directly connected next hop via IGP. The matching route with the direct next hop is called the "recursive route". The process of finding a recursive route is route recursion.

The system supports BGP load balancing based on route recursion. If multiple recursive routes to the same destination are load balanced (suppose three direct next hop addresses), BGP generates the same number of next hops to forward packets. BGP load balancing based on route recursion is always enabled by the system rather than configured using commands.

BGP differs from IGP in the implementation of load balancing in the following ways:

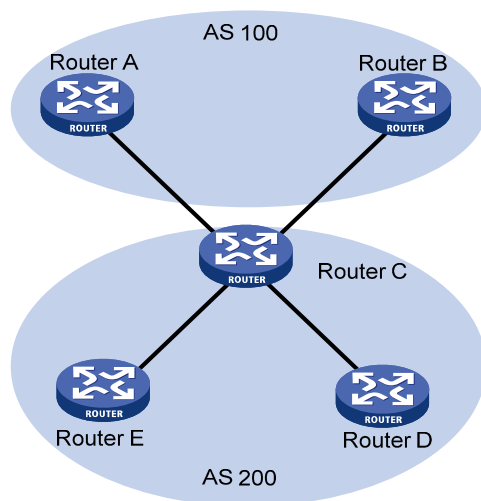
- IGP routing protocols such as RIP and OSPF compute metrics of routes, and then implement load balancing over routes with the same metric and to the same destination. The route selection criterion is metric.
- BGP has no route computation algorithm, so it cannot implement load balancing according to metrics of routes. However, BGP has abundant route selection rules, through which, it selects available routes for load balancing and adds load balancing to route selection rules.

BGP implements load balancing only on routes that have the same AS_PATH, ORIGIN, LOCAL_PREF, and MED.

BGP load balancing is applicable between EBGP peers, between IBGP peers, and between confederations.

If multiple routes to the same destination are available, BGP selects a configurable number of routes for load balancing.

Figure 80 Network diagram for BGP load balancing



In the above figure, Router D and Router E are IBGP peers of Router C. Router A and Router B both advertise a route destined for the same destination to Router C. If load balancing is configured and the two routes have the same AS_PATH attribute, ORIGIN attribute, LOCAL_PREF and MED, Router C installs both the two routes to its route table for load balancing. After that, Router C forwards to Router D and Router E the route that has AS_PATH unchanged but has NEXT_HOP changed to Router C; other BGP transitive attributes are those of the best route.

BGP route advertisement rules

The current BGP implementation supports the following route advertisement rules:

- When multiple feasible routes to a destination exist, the BGP speaker advertises only the best route to its peers.
- A BGP speaker only advertises routes that it uses.
- A BGP speaker advertises routes learned through EBGP to all BGP peers, including both EBGP and IBGP peers.
- A BGP speaker does not advertise routes from an IBGP peer to other IBGP peers.
- A BGP speaker advertises routes learned through IBGP to EBGP peers. If BGP and IGP synchronization is disabled, those routes are advertised to EBGP peers directly. If the feature is enabled, only after IGP advertises those routes, can BGP advertise the routes to EBGP peers.

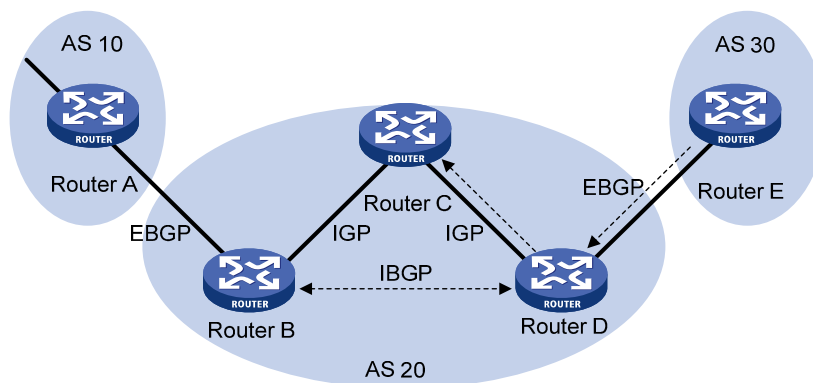
- A BGP speaker advertises all routes to a newly connected peer.

IBGP and IGP synchronization

Routing information synchronization between IBGP and IGP avoids giving wrong directions to routers outside of the local AS.

If a non-BGP router works in an AS, it can discard a packet because a destination is unreachable. As shown in Figure 81, Router E has learned a route of 8.0.0.0/8 from Router D via BGP. Router E then sends a packet to 8.0.0.0/8 through Router D, which finds from its routing table that Router B is the next hop (configured using the **peer next-hop-local** command). Because Router D has learned the route to Router B via IGP, it forwards the packet to Router C through route recursion. Router C is unaware of the route 8.0.0.0/8, so it discards the packet.

Figure 81 IBGP and IGP synchronization



For this example, if synchronization is enabled, and the route 8.0.0.0/24 received from Router B is available in its IGP routing table, Router D adds the route into its BGP routing table and advertises the route to the EBGP peer.

You can disable the synchronization feature in the following situations:

- The local AS is not a transitive AS (AS20 is a transitive AS in the above figure).
- Routers in the local AS are IBGP fully meshed.

Settlements for problems in large scale BGP networks

Route summarization

Route summarization can reduce the routing table size on a large network, and allow BGP routers to advertise only summary routes.

The system supports both manual and automatic route summarization. Manual route summarization allows you to determine the attribute of a summary route and whether to advertise the route.

Route dampening

BGP route dampening solves the issue of route instability such as route flaps—a route comes up and disappears in the routing table frequently.

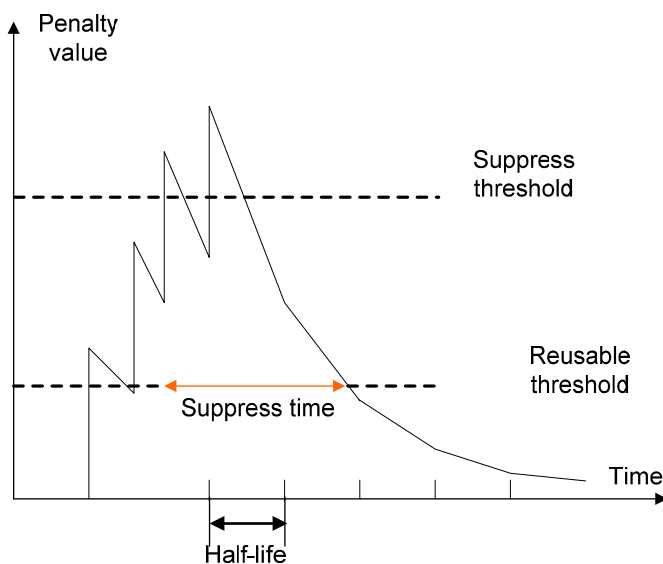
When a route flap occurs, the routing protocol sends an update to its neighbor, and then the neighbor must recalculate routes and modify the routing table. Frequent route flaps consume large bandwidth and CPU resources, which could affect network operation.

In most cases, BGP is used in complex networks, where route changes are more frequent. To solve the problem caused by route flaps, BGP route dampening is used to suppress unstable routes.

BGP route dampening, as shown in Figure 82, uses a penalty value to judge the stability of a route. The bigger the value, the less stable the route. Each time a route flap occurs, BGP adds a penalty value (1000, which is a fixed number and cannot be changed) to the route. When the penalty value of the route exceeds the suppress value, the route is suppressed from being added into the routing table or being advertised to other BGP peers.

The penalty value of the suppressed route will decrease to half of the suppress value after a period of time. This period is called "Half-life". When the value decreases to the reusable threshold value, the route is added into the routing table and advertised to other BGP peers.

Figure 82 BGP route dampening



Peer group

You can organize BGP peers with the same attributes into a group to simplify their configurations.

When a peer joins the peer group, the peer obtains the same configuration as the peer group. If the configuration of the peer group is changed, the configuration of group members is changed.

When a peer is added into a peer group, the peer has the same route update policy as the peer group to improve route distribution efficiency.

If an option is configured for both a peer and its peer group, the last configuration takes effect.

Community

A peer group provides each peer with the same policy. A community provides a group of BGP routers in several ASs with the same policy. Community is a path attribute advertised between BGP peers without being limited by AS.

A BGP router can modify the community attribute for a route before sending it to other peers.

Besides using well-known community attributes, you can define extended community attributes by using a community list to define a routing policy.

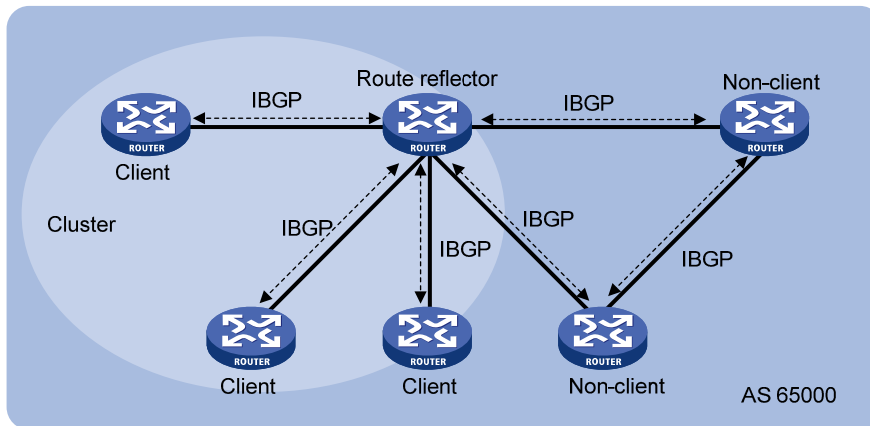
Route reflector

IBGP peers must be fully meshed to maintain connectivity. If n routers exist in an AS, the number of IBGP connections is $n(n-1)/2$, and large amounts of network and CPU resources are consumed.

Using route reflectors can resolve this issue. In an AS, a router acts as a route reflector, and other routers act as clients connecting to the route reflector. The route reflector forwards routing information between clients, so BGP sessions between clients need not be established.

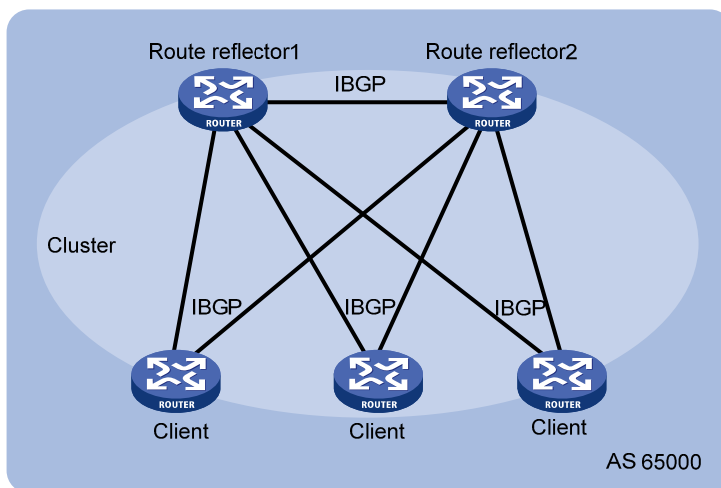
A router that is neither a route reflector nor a client is a non-client, which, as shown in [Figure 83](#), must establish BGP sessions to the route reflector and other non-clients.

Figure 83 Network diagram for a route reflector



The route reflector and clients form a cluster. In some cases, you can configure more than one route reflector in a cluster to improve network reliability and prevent a single point of failure, as shown in the following figure. The configured route reflectors must have the same Cluster_ID in order to avoid routing loops.

Figure 84 Network diagram for route reflectors



When the BGP routers in an AS are fully meshed, route reflection is unnecessary because it consumes more bandwidth resources. You can use related commands to disable route reflection.

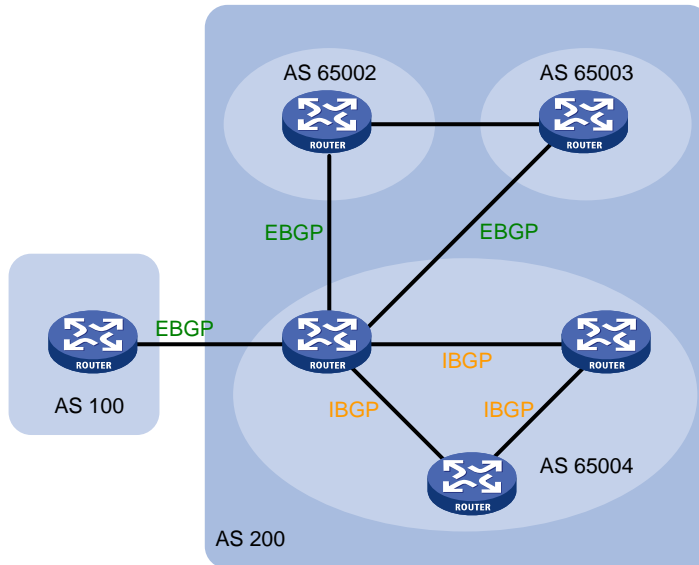
NOTE:

After route reflection is disabled between clients, routes can still be reflected between a client and a non-client.

Confederation

Confederation is another method to manage growing IBGP connections in ASs. This method splits an AS into multiple sub-ASs. In each sub-AS, IBGP peers are fully meshed, and, as shown in [Figure 85](#), intra-confederation EBGP connections are established between sub-ASs.

Figure 85 Confederation network diagram



A non-confederation BGP speaker is not required to know sub-ASs in the confederation. The ID of the confederation is the number of the AS. In the above figure, AS 200 is the confederation ID.

The deficiency of confederation is as follows:

- When changing an AS into a confederation, you must reconfigure your routers.
- The topology is changed.

In large-scale BGP networks, both route reflector and confederation can be used.

BGP GR

Graceful Restart (GR) ensures the continuity of packet forwarding when BGP restarts or an active/standby switchover occurs:

- **GR Restarter**—Graceful restarting router. It must be GR capable.
- **GR Helper**—A neighbor of the GR Restarter. It helps the GR Restarter to complete the GR process.

The following describes the BGP routing convergence process:

1. To establish a BGP session with a peer, a BGP GR Restarter sends an Open message with GR capability to the peer.
2. Upon receipt of this message, the peer is aware that the sending router is capable of Graceful Restart, and sends an Open message with GR Capability to the GR Restarter to establish a GR

session. If neither party has the GR capability, the session established between them will not be GR capable.

3. When an active/standby switchover occurs on the GR Restarter, sessions on it will go down. Then, GR-capable peers will mark all routes associated with the GR Restarter as stale. However, during the configured GR Time, they still use these routes for packet forwarding.
4. After the restart is completed, the GR Restarter will reestablish GR sessions with its peers and send a new GR message, notifying the completion of restart. Routing information is exchanged between them for the GR Restarter to create a new routing table and forwarding table and have stale routing information removed. Then the BGP routing convergence is complete.

MP-BGP

Overview

BGP-4 supports IPv4 unicasts, but does not support other network layer protocols, such as IPv6.

To support more network layer protocols, IETF extended BGP-4 by introducing Multiprotocol Extensions for BGP-4 (MP-BGP) in RFC 4760.

Routers supporting MP-BGP can communicate with routers not supporting MP-BGP.

MP-BGP extended attributes

In BGP-4, the attributes for IPv4 address format are NLRI, NEXT_HOP and AGGREGATOR (AGGREGATOR contains the IP address of the speaker generating the summary route). They are all carried in updates.

To support multiple network layer protocols, BGP-4 puts information about network layer into NLRI and NEXT_HOP. MP-BGP introduces the following path attributes:

- **MP_REACH_NLRI**—Multiprotocol Reachable NLRI, for advertising feasible routes and next hops
- **MP_UNREACH_NLRI**—Multiprotocol Unreachable NLRI, for withdrawing unfeasible routes

These path attributes are both optional non-transitive, so BGP speakers not supporting multiprotocol ignore these attributes and do not forward them to its peers.

Address family

MP-BGP uses address families to differentiate network layer protocols. For address family values, see RFC 1700, *Assigned Numbers*. The system supports multiple MP-BGP extensions, including VPN extension, and IPv6 extension. Different extensions are configured in respective address family view.

NOTE:

This chapter provides no detailed commands related to any specific extension application in MP-BGP address family view.

Protocols and standards

- RFC 1771, *A Border Gateway Protocol 4 (BGP-4)*
- RFC 2858, *Multiprotocol Extensions for BGP-4*
- RFC 3392, *Capabilities Advertisement with BGP-4*
- RFC 2918, *Route Refresh Capability for BGP-4*
- RFC 2439, *BGP Route Flap Damping*

- RFC 1997, *BGP Communities Attribute*
- RFC 2796, *BGP Route Reflection*
- RFC 3065, *Autonomous System Confederations for BGP*
- RFC 4271, *A Border Gateway Protocol 4 (BGP-4)*
- RFC 5291, *Outbound Route Filtering Capability for BGP-4*
- RFC 5292, *Address-Prefix-Based Outbound Route Filter for BGP-4*
- draft-ietf-idr-restart-08, *Graceful Restart Mechanism for BGP*

BGP configuration task list

Task	Remarks	
Configuring BGP basic functions	Creating a BGP connection	Required.
	Specifying the source interface for TCP connections	Optional.
	Allowing establishment of EBGP connection to an indirectly connected peer or peer group	Optional.
Controlling route generation	Injecting a local network	Required.
	Configuring BGP route redistribution	Use at least one approach.
	Enabling default route redistribution into BGP	Optional.
Controlling route distribution and reception	Configuring BGP route summarization	Optional.
	Advertising a default route to a peer or peer group	
	Configuring BGP route distribution/reception filtering policies	
	Enabling BGP and IGP route synchronization	
	Limiting prefixes received from a peer or peer group	
	Configuring BGP route dampening	
Configuring BGP route attributes	Configuring a shortcut route	Optional.
	Specifying a preferred value for routes received	
	Configuring preferences for BGP routes	
	Configuring the default local preference	
	Configuring the MED attribute	
	Configuring the next hop attribute	
Tuning and optimizing BGP networks	Configuring the AS-PATH attribute	Optional.
	Configuring the BGP keepalive interval and holdtime	Optional.

Task	Remarks	
Configuring the interval for sending the same update	Optional.	
Configuring BGP soft-reset	Optional.	
Enabling the BGP ORF capability	Optional.	
Enabling 4-byte AS number suppression	Optional.	
Setting the DSCP value for BGP packets	Optional.	
Enabling quick EBGP session reestablishment	Optional.	
Enabling MD5 authentication for TCP connections	Optional.	
Configuring BGP load balancing	Optional.	
Forbidding session establishment with a peer or peer group	Optional.	
Configuring a large scale BGP network	Configuring BGP peer groups	Optional.
	Configuring BGP community	Optional.
	Configuring a BGP route reflector	Optional.
	Configuring a BGP confederation	Optional.
Configuring BGP GR	Optional.	
Enabling trap	Optional.	
Enabling logging of peer state changes	Optional.	
Configuring BFD for BGP	Optional.	

Configuring BGP basic functions

This section does not differentiate between BGP and MP-BGP.

Configuration prerequisites

The neighboring nodes are accessible to each other at the network layer.

Creating a BGP connection

A router ID is the unique identifier of a BGP router in an AS.

- To ensure the uniqueness of a router ID and enhance network reliability, you can specify in BGP view the IP address of a local loopback interface as the router ID.
- If no router ID is specified in BGP view, the global router ID is used.
- If the global router ID is used and then it is removed, the system will select a new router ID.
- If the router ID is specified in BGP view, using the **undo router-id** command can make the system select a new router ID.

To create a BGP connection:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure a global router ID.	router id <i>router-id</i>	Optional. Not configured by default. If no global router ID is configured, the highest loopback interface IP address—if any—is used as the router ID. If no loopback interface IP address is available, the highest physical interface IP address is used, regardless of the interface status.
3. Enable BGP and enter BGP view.	bgp <i>as-number</i>	Not enabled by default.
4. Specify a router ID.	router-id <i>router-id</i>	Optional. By default, the global router ID is used.
5. Specify a peer or a peer group and its AS number.	peer { <i>group-name</i> <i>ip-address</i> } as-number <i>as-number</i>	Not specified by default.
6. Enable the default use of IPv4 unicast address family for the peers that are established using the peer as-number command.	default ipv4-unicast	Optional. Enabled by default.
7. Enable a peer.	peer <i>ip-address</i> enable	Optional. Enabled by default. A router can reside in only one AS, so the router can run only one BGP process.
8. Configure a description for a peer/peer group.	peer { <i>group-name</i> <i>ip-address</i> } description <i>description-text</i>	Not configured by default. You must create a peer group before configuring it.

Specifying the source interface for TCP connections

BGP uses TCP as the transport layer protocol. By default, BGP uses the output interface of the optimal route to a peer/peer group as the source interface for establishing TCP connections to the peer or peer group.

If a BGP router has multiple links to a peer, and the source interface fails, BGP must reestablish TCP connections, causing network oscillation. To enhance stability of BGP connections, HP recommends using a loopback interface as the source interface.

To establish multiple BGP connections between two routers, you must specify the source interface for establishing TCP connections to each peer on the local router; otherwise, the local BGP router may fail to establish TCP connections to a peer when using the outbound interface of the best route to the peer as the source interface.

To specify the source interface for TCP connections:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Specify the source interface for establishing TCP connections to a peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> } connect-interface <i>interface-type</i> <i>interface-number</i>	By default, BGP uses the outbound interface of the best route to the BGP peer or peer group as the source interface for establishing a TCP connection to the peer or peer group.

Allowing establishment of EBGP connection to an indirectly connected peer or peer group

Direct physical links must be available between EBGP peers. If they are not, use the **peer ebgp-max-hop** command to establish a TCP connection over multiple hops between two peers.

Do not configure the **peer ebgp-max-hop** command if the two EBGP peers are directly connected.

To allow establishment of EBGP connection to an indirectly connected peer or peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Allow the establishment of EBGP connection to an indirectly connected peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> } ebgp-max-hop [<i>hop-count</i>]	Optional. Not allowed by default.

Controlling route generation

Different from IGP, BGP focuses on route generation and advertisement control and optimal route selection.

Generating BGP routes can be done in the following ways:

- Configure BGP to advertise local networks.
- Configure BGP to redistribute routes from other routing protocols, including the default route.

Configuration prerequisites

BGP connections must be created.

Injecting a local network

In BGP view, you can inject a local network to allow BGP to advertise to BGP peers. The origin attribute of routes advertised in this way is IGP. You can also reference a routing policy to flexibly control route advertisement. The network to be injected must be available in the local IP routing table.

To inject a local network:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Inject a network to the BGP routing table.	network <i>ip-address</i> [<i>mask</i> <i>mask-length</i>] route-policy <i>route-policy-name</i>	Optional. Not injected by default.

Configuring BGP route redistribution

BGP does not find routes by itself. Rather, it redistributes routing information in the local AS from other routing protocols. During route redistribution, you can configure BGP to filter routing information from specific routing protocols.

The origin attribute of routes redistributed using the **import-route** command is INCOMPLETE.

Only active routes can be redistributed. You can use the **display ip routing-table protocol** command to display route state information.

To configure BGP route redistribution:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enable route redistribution from a routing protocol into BGP.	import-route <i>protocol</i> [{ <i>process-id</i> all-processes } [allow-direct med <i>med-value</i> route-policy <i>route-policy-name</i>] *]	Not enabled by default. The allow-direct keyword is available only when the specified routing protocol is OSPF.

Enabling default route redistribution into BGP

Using the **import-route** command cannot redistribute a default route.

To enable default route redistribution into BGP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enable route redistribution from a routing protocol into BGP.	import-route <i>protocol</i> [{ <i>process-id</i> all-processes } [allow-direct med <i>med-value</i> route-policy <i>route-policy-name</i>] *]	Not redistributed by default. The allow-direct keyword is available only when the specified routing protocol is OSPF.
4. Enable default route redistribution into BGP.	default-route imported	Optional. Not enabled by default.

Controlling route distribution and reception

Configuration prerequisites

BGP connections must be created.

Configuring BGP route summarization

To reduce the routing table size on medium and large BGP networks, you need to configure route summarization on BGP routers. BGP supports automatic and manual summarization modes. Manual summary routes have a higher priority than automatic ones.

Configuring automatic route summarization

After automatic route summarization is configured, BGP summarizes redistributed IGP subnets to advertise only natural networks. Routes injected with the **network** command cannot be summarized.

To configure automatic route summarization:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Configure automatic route summarization.	summary automatic	Not configured by default.

Configuring manual route summarization

By configuring manual route summarization, you can summarize both redistributed routes and routes injected using the **network** command and determine the mask length for a summary route as needed.

To configure BGP manual route summarization:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Configure manual route summarization.	aggregate <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } [as-set attribute-policy <i>route-policy-name</i> detail-suppressed origin-policy <i>route-policy-name</i> suppress-policy <i>route-policy-name</i>]*	Not configured by default.

Advertising a default route to a peer or peer group

After this task is configured, the BGP router sends a default route with the next hop being itself to the specified peer or peer group. This action is taken regardless of whether the default route is available in the routing table.

To advertise a default route to a peer or peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Advertise a default route to a peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> } default-route-advertise [route-policy <i>route-policy-name</i>]	Not advertised by default.

Configuring BGP route distribution/reception filtering policies

Configuration prerequisites

Configure following filters as needed:

- ACL
- IP prefix list
- Routing policy
- AS-path ACL

For how to configure an ACL, see *ACL and QoS Configuration Guide*.

For how to configure an IP prefix list, routing policy, and AS-path ACL, see "[Configuring routing policies](#)."

Configuring BGP route distribution filtering policies

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A

Step	Command	Remarks
3. Configure BGP route distribution filtering policies.	<ul style="list-style-type: none"> Configure the filtering of redistributed routes: filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } export [direct isis <i>process-id</i> ospf <i>process-id</i> rip <i>process-id</i> static] Reference a routing policy to filter advertisements to a peer or peer group: peer { <i>group-name</i> <i>ip-address</i> } route-policy <i>route-policy-name</i> export Reference an ACL to filter advertisements to a peer or peer group: peer { <i>group-name</i> <i>ip-address</i> } filter-policy <i>acl-number</i> export Reference an AS path ACL to filter routing information sent to a peer or peer group: peer { <i>group-name</i> <i>ip-address</i> } as-path-acl <i>as-path-acl-number</i> export Reference an IP prefix list to filter routing information sent to a peer or peer group: peer { <i>group-name</i> <i>ip-address</i> } ip-prefix <i>ip-prefix-name</i> export 	<p>Configure at least one command. Not configured by default. You can configure a filtering policy as needed.</p> <p>If several filtering policies are configured, they are applied in the following sequence:</p> <ol style="list-style-type: none"> filter-policy export peer filter-policy export peer as-path-acl export peer ip-prefix export peer route-policy export <p>Only routes passing the first policy can go to the next, and only routes passing all the configured policies can be advertised.</p>

Configuring BGP route reception filtering policies

Only routes permitted by the configured filtering policies can be installed into the local BGP routing table. The members of a peer group can have different route reception filtering policies from the peer group.

To configure BGP route reception filtering policies:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A

Step	Command	Remarks
3. Configure BGP route reception filtering policies.	<ul style="list-style-type: none"> Filter incoming routes with an ACL or IP prefix list: filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } import 	Configure at least one command. No route reception filtering is configured by default. If several filtering policies are configured, they are applied in the following sequence: 4. filter-policy import 5. peer filter-policy import 6. peer as-path-acl import 7. peer ip-prefix import 8. peer route-policy import Only routes passing all the configured policies can be received.
	<ul style="list-style-type: none"> Reference a routing policy to filter routes from a peer or peer group: peer { <i>group-name</i> <i>ip-address</i> } route-policy <i>route-policy-name</i> import 	
	<ul style="list-style-type: none"> Reference an ACL to filter routing information from a peer or peer group: peer { <i>group-name</i> <i>ip-address</i> } filter-policy <i>acl-number</i> import 	
	<ul style="list-style-type: none"> Reference an AS path ACL to filter routing information from a peer or peer group: peer { <i>group-name</i> <i>ip-address</i> } as-path-acl <i>as-path-acl-number</i> import 	
	<ul style="list-style-type: none"> Reference an IP prefix list to filter routing information from a peer or peer group: peer { <i>group-name</i> <i>ip-address</i> } ip-prefix <i>ip-prefix-name</i> import 	

Enabling BGP and IGP route synchronization

By default, upon receiving an IBGP route, a BGP router checks the route's next hop. If the next hop is reachable, the BGP router advertises the route to EBGP peers. If the synchronization feature is configured, in addition to the reachability check of the next hop, the BGP router must find an active IGP route with the same destination network segment before it can advertise the IBGP route (use the **display ip routing-table protocol** command to check the IGP route state).

To enable BGP and IGP synchronization:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enable synchronization between BGP and IGP.	synchronization	Not enabled by default.

Limiting prefixes received from a peer or peer group

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A

Step	Command	Remarks
3. Configure the maximum number of prefixes allowed to be received from a peer or peer group.	<ul style="list-style-type: none"> Specify the maximum number of prefixes that can be received from a peer or peer group: peer { <i>group-name</i> <i>ip-address</i> } route-limit <i>prefix-number</i> [<i>percentage-value</i>] Specify the maximum number of prefixes that can be received from a peer or peer group: peer { <i>group-name</i> <i>ip-address</i> } route-limit <i>prefix-number</i> alert-only [<i>percentage-value</i>] Specify the maximum number of prefixes that can be received from a peer or peer group: peer { <i>group-name</i> <i>ip-address</i> } route-limit <i>prefix-number</i> reconnect <i>reconnect-time</i> [<i>percentage-value</i>] 	<p>Use one of the commands.</p> <p>No limit is configured by default.</p> <p>If the specified maximum number is reached:</p> <ul style="list-style-type: none"> For approach 1, the router breaks down the BGP connection to the peer. For approach 2, the router outputs alert information but does not break down the BGP connection to the peer. For approach 3, the router breaks down the BGP connection to the peer and then reestablishes a BGP connection to the peer.

Configuring BGP route dampening

By configuring BGP route dampening, you can suppress unstable routes from being added to the local routing table or being advertised to BGP peers.

To configure BGP route dampening:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Configure BGP route dampening.	dampening [<i>half-life-reachable</i> <i>half-life-unreachable</i> <i>reuse</i> <i>suppress ceiling</i> route-policy <i>route-policy-name</i>] *	Not configured by default.

Configuring a shortcut route

An EBGp route received has a priority of 255, which is lower than a local route. This task allows you to configure an EBGp route as a shortcut route. It has the same priority as a local route and is more likely to become the optimal route.

To configure a shortcut route:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A

Step	Command	Remarks
3. Configure a shortcut route.	network <i>ip-address</i> [<i>mask</i> <i>mask-length</i>] short-cut	Optional. By default, an EBGP route received has a priority of 255.

Configuring BGP route attributes

Configuration prerequisites

BGP connections must be created.

Specifying a preferred value for routes received

By default, routes received from a peer have a preferred value of 0. Among multiple routes that have the same destination and mask and are learned from different peers, the one with the greatest preferred value is selected as the route to the destination.

To specify a preferred value for routes from a peer or peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Specify a preferred value for routes received from a peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> } preferred-value <i>value</i>	Optional. By default, the preferred value is 0.

Configuring preferences for BGP routes

A router can run multiple routing protocols with each having a preference. If they find the same route, the route found by the routing protocol with the highest preference is selected.

This task allows you to configure preferences for external, internal, and local BGP routes, and reference a routing policy to set preferences for matching routes as needed.

To configure preferences for BGP routes:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Configure preferences for external, internal, and local BGP routes.	preference { <i>external-preference</i> <i>internal-preference</i> <i>local-preference</i> route-policy <i>route-policy-name</i> }	Optional. The default preferences of external, internal, and local BGP routes are 255, 255, and 130.

Configuring the default local preference

The local preference is used to determine the best route for traffic leaving the local AS. When a BGP router obtains from several IBGP peers multiple routes to the same destination but with different next hops, it considers the route with the highest local preference as the best route.

This task allows you to specify the default local preference for routes sent to IBGP peers.

To specify the default local preference:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Configure the default local preference.	default local-preference <i>value</i>	Optional. 100 by default.

Configuring the MED attribute

MED is used to determine the best route for traffic going into an AS. When a BGP router obtains from EBGP peers multiple routes to the same destination but with different next hops, it considers the route with the smallest MED value as the best route if other conditions are the same.

Configuring the default MED value

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Configure the default MED value.	default med <i>med-value</i>	Optional. 0 by default.

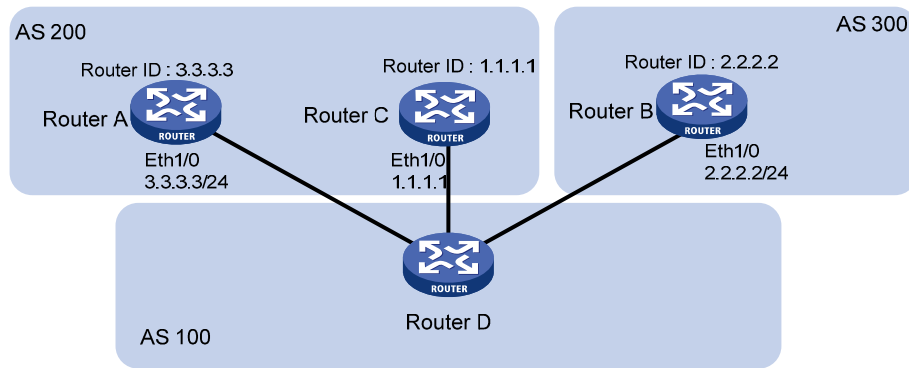
Enabling the comparison of MED of routes from different ASs

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enable the comparison of MED of routes from different ASs.	compare-different-as-med	Not enabled by default.

Enabling the comparison of MED of routes from each AS

Route learning sequence may affect optimal route selection.

Figure 86 Route selection based on MED



As shown in Figure 86, Router D learns network 10.0.0.0 from both Router A and Router B. Because Router B has a smaller router ID, the route learned from it is optimal.

Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>i 10.0.0.0	2.2.2.2	50		0	300e
* i	3.3.3.3	50		0	200e

When Router D learns network 10.0.0.0 from Router C, it compares the route with the optimal route in its routing table. Because Router C and Router B reside in different ASs, BGP will not compare the MEDs of the two routes. Router C has a smaller router ID than Router B, the route from Router C becomes optimal.

Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>i 10.0.0.0	1.1.1.1	60		0	200e
* i 10.0.0.0	2.2.2.2	50		0	300e
* i	3.3.3.3	50		0	200e

However, Router C and Router A reside in the same AS, and Router C has a greater MED, so network 10.0.0.0 learned from Router C cannot be optimal.

You can configure the **bestroute compare-med** command on Router D. After that, Router D puts routes received from the same AS into a group. Router D then selects the route with the lowest MED from the same group, and compares routes from different groups. This mechanism avoids the above-mentioned problem. The following output is the BGP routing table on Router D after the comparison of MED of routes from each AS is enabled. Network 10.0.0.0 learned from Router B is the optimal route.

Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>i 10.0.0.0	2.2.2.2	50		0	300e
* i 10.0.0.0	3.3.3.3	50		0	200e
* i	1.1.1.1	60		0	200e

BGP load balancing cannot be implemented because load balanced routes must have the same AS-path attribute.

To enable the comparison of MED of routes from each AS:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A
3. Enable the comparison of MED of routes from each AS.	bestroute compare-med	Optional. Not enabled by default.

Enabling the comparison of MED of routes from confederation peers

The MED attributes of routes from confederation peers are not compared if their AS-path attributes contain AS numbers that do not belong to the confederation, such as these three routes: AS-path attributes of them are 65006 65009, 65007 65009, and 65008 65009; and MED values of them are 2, 3, and 1. Because the third route contains an AS number that does not belong to the confederation, the first route becomes the optimal route.

To enable the comparison of MED of routes from confederation peers:

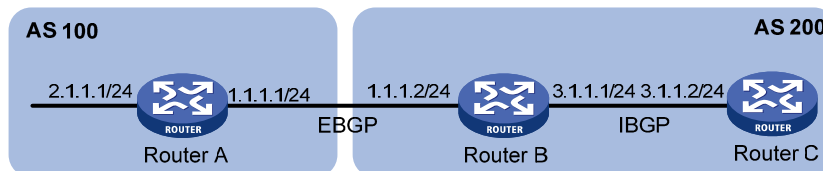
Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A
3. Enable the comparison of MED of routes from confederation peers.	bestroute med-confederation	Optional. Not enabled by default.

Configuring the next hop attribute

By default, when advertising routes to an IBGP peer or peer group, a BGP router does not set itself as the next hop. However, to ensure a BGP peer can find the correct next hop in some cases, you must configure the router as the next hop for routes sent to the peer.

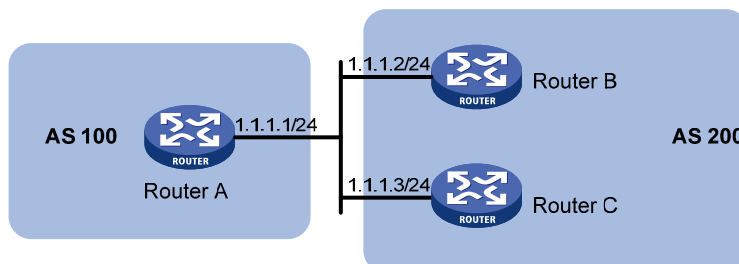
For example, as shown in [Figure 87](#), Router A and Router B establish an EBGP neighbor relationship, and Router B and Router C establish an IBGP neighbor relationship. When Router B advertises a network learned from Router A to Router C, if Router C has no route to IP address 1.1.1.1/24, you must configure Router B to set itself as the next hop (3.1.1.1/24) for the route to be sent to Router C.

Figure 87 Next hop attribute configuration 1



If a BGP router has two peers on a common broadcast network, it does not set itself as the next hop for routes sent to an EBGP peer by default. As shown in [Figure 88](#), Router A and Router B establish an EBGP neighbor relationship, and Router B and Router C establish an IBGP neighbor relationship. They are on the same broadcast network 1.1.1.0/24. When Router B sends EBGP routes to Router A, it does not set itself as the next hop by default. However, you can configure Router B to set it as the next hop (1.1.1.2/24) for routes sent to Router A by using the **peer next-hop-local** command as needed.

Figure 88 Next hop attribute configuration 2



If you have configured BGP load balancing on a BGP router, the router will set it as the next hop for routes sent to an IBGP peer or peer group. This is done regardless of whether the **peer next-hop-local** command is configured.

To configure the next hop attribute:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Specify the router as the next hop of routes sent to a peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> } next-hop-local	Optional. By default, the router sets it as the next hop for routes sent to an EBGP peer or peer group, but does not set it as the next hop for routes sent to an IBGP peer or peer group.

Configuring the AS-PATH attribute

Permitting local AS number to appear in routes from a peer or peer group

BGP checks whether the AS_PATH attribute of a route from a peer contains the local AS number. If so, it discards the route to avoid routing loops.

To permit local AS number to appear in routes from a peer or peer group and specify the appearance times.

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Permit local AS number to appear in routes from a peer or peer group and specify the appearance times.	peer { <i>group-name</i> <i>ip-address</i> } allow-as-loop [<i>number</i>]	Optional. By default, the local AS number is not allowed.

Disabling BGP from considering AS_PATH during best route selection

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Disable BGP from considering AS_PATH during best route selection.	bestroute as-path-neglect	Optional. By default, BGP considers AS_PATH during best route selection.

Specifying a fake AS number for a peer or peer group

When Router A in AS 2 is moved to AS 3, you can configure Router A to specify a fake AS number of 2 for created connections to EBGP peers or peer groups. In this way, these EBGP peers still think Router A is in AS 2 and need not change their configurations. This feature ensures uninterrupted BGP services.

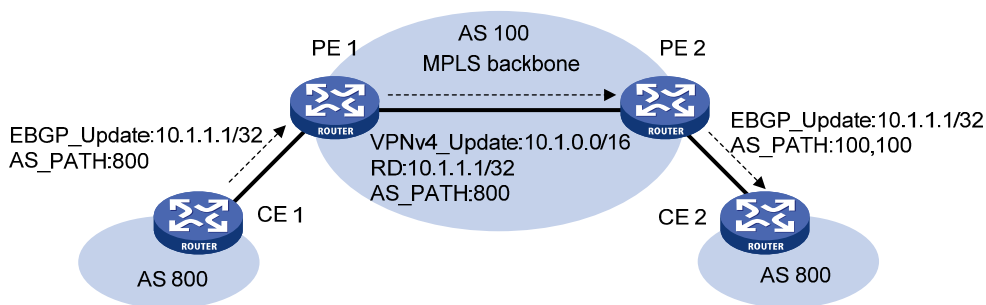
To specify a fake AS number for a peer or peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A
3. Specify a fake AS number for a peer or peer group.	peer { group-name ip-address } fake-as as-number	Optional. Not specified by default. This command is only applicable to an EBGP peer or peer group.

Configuring AS number substitution

In L3VPN, if EBGP is used between PE and CE, sites in different geographical areas must have different AS numbers assigned to ensure correct route advertisement. If different CEs use the same AS number, you must configure the relevant PE to replace the AS number of the CE as its own AS number. This feature is used for route advertisement only.

Figure 89 AS number substitution configuration



As shown in the above figure, CE 1 and CE 2 use the same AS number of 800. If AS number substitution for CE 2 is configured on PE 2, and PE 2 receives a BGP update sent from CE 1, PE 2 replaces AS number 800 as its own AS number 100. Similar configuration must also be made on PE 1.

To configure AS number substitution for a peer or peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A
3. Replace the AS number of a peer or peer group in the AS_PATH attribute as the local AS number.	peer { group-name ip-address } substitute-as	Not configured by default. Improper AS number substitution configuration may cause route loops; use this command with caution.

Removing private AS numbers from updates to a peer or peer group

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Configure BGP to remove private AS numbers from the AS_PATH attribute of updates to a peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> } public-as-only	By default, BGP updates carry private AS numbers.

Ignoring the first AS number of EBGP route updates

Typically, BGP checks the AS_PATH attribute of a route update received from a peer. If the first AS number is not that of the BGP peer, the BGP router discards the route update.

To ignore the first AS number of EBGP route updates:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Configure BGP to ignore the first AS number of EBGP route updates.	ignore-first-as	By default, BGP checks the first AS number of EBGP route updates.

Tuning and optimizing BGP networks

Configuration prerequisites

BGP connections must be created.

Configuring the BGP keepalive interval and holdtime

After establishing a BGP connection, two routers send keepalive messages periodically to each other to keep the connection. If a router receives no keepalive or update message from the peer within the holdtime, it tears down the connection.

You can configure the keepalive interval and holdtime globally or for a specific peer or peer group. The actual keepalive interval and holdtime depend on the following cases:

- If the holdtime settings on the local and peer routers are different, the smaller one is used. The holdtime is no less than three seconds unless it is set to 0.
- If the keepalive interval is 0 and the negotiated holdtime is not 0, the actual keepalive interval equals one-third of the holdtime. The maximum keepalive interval must be one third of the holdtime and no less than one second.
- If the keepalive interval is not 0, the actual keepalive interval is the smaller one between one third of the holdtime and the keepalive interval.

Follow these guidelines when you configure BGP keepalive interval and holdtime:

- The intervals set with the **peer timer** command are preferred to those set with the **timer** command.

- If the router has established a neighbor relationship with a peer, you must reset the BGP connection to validate the new set timers.
- The **timer** command takes effect for only new connections.
- After **peer timer** command is executed, the peer connection is closed at once, and a new connection to the peer is negotiated using the configured hold time.

To configure BGP keepalive interval and holdtime:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Configure BGP keepalive interval and holdtime.	<ul style="list-style-type: none"> • Configure the global keepalive interval and holdtime: timer keepalive <i>keepalive</i> hold <i>holdtime</i> • Configure the keepalive interval and holdtime for a peer or peer group: peer { <i>group-name</i> <i>ip-address</i> } timer keepalive <i>keepalive</i> hold <i>holdtime</i> 	<p>Optional.</p> <p>By default, the keepalive interval is 60 seconds, and holdtime is 180 seconds.</p>

Configuring the interval for sending the same update

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Configure the interval for sending the same update to a peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> } route-update-interval <i>interval</i>	<p>Optional.</p> <p>The intervals for sending the same update to an IBGP peer and an EBGP peer default to 15 seconds and 30 seconds.</p>

Configuring BGP soft-reset

After modifying the route selection policy, reset BGP connections to make the new one take effect.

The current BGP implementation supports the route-refresh feature that enables dynamic route refresh without tearing down BGP connections.

However, if a peer not supporting route-refresh exists in the network, you must configure the **peer keep-all-routes** command to save all routes from the peer, which are used during applying the new route selection policy.

Configuring automatic soft-reset

After route refresh is enabled for peers and a policy is modified, the router advertises a route-refresh message to the peers, which then resend their routing information to the router. After receiving the routing information, the router performs dynamic route update by using the new policy.

To enable BGP route refresh for a peer or peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enable BGP route refresh for a peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> } capability-advertise route-refresh	Optional. Enabled by default.

Configuring manual soft-reset

If a BGP peer does not support route-refresh, you must save updates from the peer on the local router by using the **peer keep-all-routes** command, and use the **refresh bgp** command to refresh the BGP routing table.

If the BGP peer does not support route-refresh and the **peer keep-all-routes** command is not configured for it, you need to decide whether to manually disconnect the peer to learn routes again according to the impact of the new policy.

Following these steps to save all route updates from a peer or peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Disable BGP route-refresh and multi-protocol extension capability for a peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> } capability-advertise conventional	Enabled by default.
4. Save all routes from a peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> } keep-all-routes	Not saved by default.
5. Return to user view.	return	N/A
6. Perform manual soft reset on BGP connections.	refresh bgp { all <i>ip-address</i> group <i>group-name</i> external internal } { export import }	N/A

Enabling the BGP ORF capability

The BGP Outbound Route Filtering (ORF) feature allows a BGP speaker to send its BGP peer a set of ORFs through route-refresh messages. The peer then applies the ORFs, in addition to its local routing policies (if any), to filter updates to the BGP speaker, reducing the number of exchanged Update messages and saving network resources.

After you enable the BGP ORF capability, the local BGP router negotiates the ORF capability with the BGP peer through Open messages (determines whether to carry ORF information in messages, and if yes,

whether to carry non-standard ORF information in the packets). After completing the negotiation process and establishing the neighboring relationship, the BGP router and its BGP peer can exchange ORF information through specific route-refresh messages.

For the parameters configured on both sides for ORF capability negotiation, see [Table 8](#).

To enable the BGP ORF capability:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A
3. Enable BGP route refresh for a peer or peer group.	peer { group-name ip-address } capability-advertise route-refresh	Enabled by default.
4. Enable the non-standard ORF capability for a BGP peer or peer group.	peer { group-name ip-address } capability-advertise orf non-standard	Optional. By default, standard BGP ORF capability defined in RFC 5291 and RFC 5292 is supported. If the peer supports only non-standard ORF, you need to configure this command.
5. Enable the ORF capability for a BGP peer or peer group.	peer { group-name ip-address } capability-advertise orf ip-prefix { both receive send }	Disabled by default.

Table 8 Description of the both, send, and receive parameters and the negotiation result

Local parameter	Peer parameter	Negotiation result
send	<ul style="list-style-type: none"> • receive • both 	The ORF sending capability is enabled locally and the ORF receiving capability is enabled on the peer.
receive	<ul style="list-style-type: none"> • send • both 	The ORF receiving capability is enabled locally and the ORF sending capability is enabled on the peer.
both	both	Both the ORF sending and receiving capabilities are enabled locally and on the peer.

Enabling 4-byte AS number suppression

When a switch that supports 4-byte AS numbers sends an Open message for peer relationship establishment, the Optional parameters field of the message indicates that the AS number occupies four bytes—in the range of 1 to 4294967295. If the peer device does not support 4-byte AS numbers (for examples, it supports only 2-byte AS numbers), the peer relationship cannot be established.

After you enable the 4-byte AS number suppression function, the peer device can then process the Open message even though it does not support 4-byte AS numbers, and the BGP peer relationship can be established.

If the peer device supports 4-byte AS numbers, do not enable the 4-byte AS number suppression function; otherwise, the BGP peer relationship cannot be established.

To enable 4-byte AS number suppression:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enable 4-byte AS number suppression.	peer { <i>group-name</i> <i>ip-address</i> } capability-advertise suppress-4-byte-as	Disabled by default.

Setting the DSCP value for BGP packets

An IPv4 packet header contains an 8-bit Type of Service (TOS) field. As defined in RFC 2474, the first six bits set the Differentiated Services Code Point (DSCP) value and the last two bits are reserved. Network devices use the DSCP value as a reference to determine the packet priority for transmission.

You can set the DSCP value for BGP packets.

To configure the DSCP value for packets sent to a BGP peer or peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view or BGP-VPN view.	<ul style="list-style-type: none"> • Enter BGP view: bgp <i>as-number</i> • Enter BGP-VPN instance view: <ul style="list-style-type: none"> a. bgp <i>as-number</i> b. ipv4-family vpn-instance <i>vpn-instance-name</i> 	Use either approach.
3. Set the DSCP value for the BGP packets sent to the specified BGP peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> } dscp <i>dscp-value</i>	Optional. By default, the DSCP value in BGP packets is 48.

Enabling quick EBGP session reestablishment

If the router receives no keepalive messages from a BGP peer within the holdtime, it disconnects from the peer.

With quick EBGP connection reestablishment enabled, the router will reestablish a session to the EBGP peer immediately when the link to a directly connected EBGP peer is down.

To enable quick EBGP session reestablishment:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enable quick EBGP session reestablishment.	ebgp-interface-sensitive	Optional. Not enabled by default.

Enabling MD5 authentication for TCP connections

BGP employs TCP as the transport protocol. To increase security, configure BGP to perform MD5 authentication when establishing a TCP connection. The two parties must have the same password configured to establish TCP connections.

BGP MD5 authentication is not for BGP packets, but for TCP connections. If the authentication fails, no TCP connection can be established.

To enable MD5 authentication for TCP connections:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enable MD5 authentication when establishing a TCP connection to the peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> } password { cipher simple } <i>password</i>	Optional. Not enabled by default.

Configuring BGP load balancing

If multiple paths to a destination exist, you can configure load balancing over such paths to improve link utilization.

To configure BGP load balancing:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Configure the maximum number of BGP routes for load balancing.	balance <i>number</i>	Optional. By default, Load balancing is not enabled.

Forbidding session establishment with a peer or peer group

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Forbid session establishment with a peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> } ignore	Optional. Not forbidden by default.

Configuring a large scale BGP network

In a large-scale BGP network, configuration and maintenance become difficult because of the large numbers of BGP peers. To facilitate configuration, you can configure peer group, community, route reflector, or confederation as needed.

Configuration prerequisites

Peering nodes are accessible to each other at the network layer.

Configuring BGP peer groups

A peer group is a group of peers with the same route selection policy.

In a large-scale network, many peers may use the same route selection policy. You can configure a peer group and add these peers into this group. In this way, peers can share the same policy as the peer group. When the policy of the group is modified, the modification also applies to peers in it, simplifying configuration.

A peer group is an IBGP peer group if peers in it belong to the same AS, and is an EBGP peer group if peers in it belong to different ASs.

If a peer group has peers added, you cannot remove its AS number using the **undo** form of the command or change its AS number.

Configuring an IBGP peer group

After you create an IBGP peer group and then add a peer into it, the system creates the peer in BGP view and specifies the local AS number for the peer.

To configure an IBGP peer group:

Step	Command
1. Enter system view.	system-view
2. Enter BGP view.	bgp <i>as-number</i>
3. Create an IBGP peer group.	group <i>group-name</i> [internal]
4. Add a peer into the IBGP peer group.	peer <i>ip-address</i> group <i>group-name</i>

Configuring an EBGP peer group

If peers in an EBGP group belong to the same external AS, the EBGP peer group is a pure EBGP peer group; if not, it is a mixed EBGP peer group.

Use one of the following approaches to configure an EBGP peer group:

- Create the EBGP peer group, specify its AS number, and add peers into it. All the added peers share the same AS number.
- Create the EBGP peer group, specify an AS number for a peer, and add the peer into the peer group.
- Create the EBGP peer group and add a peer into it with an AS number specified.

To configure an EBGP peer group using Approach 1:

Step	Command
1. Enter system view.	system-view
2. Enter BGP view.	bgp <i>as-number</i>
3. Create an EBGP peer group.	group <i>group-name</i> external
4. Specify the AS number for the group.	peer <i>group-name</i> as-number <i>as-number</i>
5. Add the peer into the group.	peer <i>ip-address</i> group <i>group-name</i>

NOTE:

- Do not specify any AS number for a peer before adding it into the peer group.
- All of the added peers have the same AS number as that of the peer group.

To configure an EBGP peer group using Approach 2:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Create an EBGP peer group.	group <i>group-name</i> external	N/A
4. Specify an AS number for a peer.	peer <i>ip-address</i> as-number <i>as-number</i>	N/A
5. Add the peer into the group.	peer <i>ip-address</i> group <i>group-name</i> [as-number <i>as-number</i>]	The AS number can be either specified or not specified in the command. If specified, the AS number must be the same as that specified for the peer with the peer <i>ip-address</i> as-number <i>as-number</i> command.

NOTE:

Peers added in the group can have different AS numbers.

To configure an EBGP peer group using Approach 3:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Create an EBGP peer group.	group <i>group-name</i> external	N/A
4. Add a peer into the group and specify its AS number.	peer <i>ip-address</i> group <i>group-name</i> as-number <i>as-number</i>	N/A

NOTE:

- Do not specify any AS number for a peer before adding it into the peer group.
 - Peers added in the group can have different AS numbers.
-

Configuring BGP community

A BGP community is a group of destinations with the same characteristics. It has no geographical boundaries and is independent of ASs.

You can configure a routing policy to define which destinations belong to a BGP community and then advertise the community attribute to a peer or peer group.

You can apply a routing policy to filter routes advertised to or received from a peer or peer group according to the community attribute, which helps simplify policy configuration and management.

For how to configure a routing policy, see "[Configuring routing policies.](#)"

To configure BGP community:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Advertise the community attribute to a peer or peer group.	<ul style="list-style-type: none">• Advertise the community attribute to a peer or peer group: peer { <i>group-name</i> <i>ip-address</i> } advertise-community• Advertise the extended community attribute to a peer or peer group: peer { <i>group-name</i> <i>ip-address</i> } advertise-ext-community	Not configured by default.
4. Apply a routing policy to routes advertised to a peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> } route-policy <i>route-policy-name</i> export	Not configured by default.

Configuring a BGP route reflector

If an AS has many BGP routers, you can configure them as a cluster by configuring one of them as a route reflector and others as clients to reduce IBGP connections.

To enhance network reliability and prevent single point of failure, specify multiple route reflectors for a cluster. The route reflectors in the cluster must have the same cluster ID to avoid routing loops.

It is not required to make clients of a route reflector fully meshed. The route reflector forwards routing information between clients. If clients are fully meshed, disable route reflection between clients to reduce routing costs.

Generally, a cluster has only one route reflector, and the router ID is used to identify the cluster. You can configure multiple route reflectors to improve network stability. To avoid routing loops, specify the same cluster ID for these route reflectors by using the **reflector cluster-id** command.

To configure a BGP route reflector:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Configure the router as a route reflector and specify a peer or peer group as its client.	peer { <i>group-name</i> <i>ip-address</i> } reflect-client	Not configured by default.
4. Enable route reflection between clients.	reflect between-clients	Optional. Enabled by default.
5. Configure the cluster ID of the route reflector.	reflector cluster-id <i>cluster-id</i>	Optional. By default, a route reflector uses its router ID as the cluster ID.

Configuring a BGP confederation

Configuring a BGP confederation is another way for reducing IBGP connections in an AS.

A confederation contains sub ASs. In each sub AS, IBGP peers are fully meshed. Between sub ASs, EBGP connections are established.

If routers not compliant with RFC 3065 exist in the confederation, use the **confederation nonstandard** command to make the local router compatible with these routers.

Configuring a BGP confederation

After you split an AS into multiple sub ASs, you can configure a router in a sub AS as follows:

1. Enable BGP and specify the AS number of the router.
2. Specify the confederation ID. From an outsider's perspective, the sub ASs of the confederation is a single AS, which is identified by the confederation ID.
3. If the router must establish EBGP connections to other sub ASs, you must specify the peering sub ASs in the confederation.

A confederation contains a maximum of 32 sub ASs. The AS number of a sub AS is effective only in the confederation.

To configure a BGP confederation:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Configure a confederation ID.	confederation id <i>as-number</i>	Not configured by default.
4. Specify peering sub ASs in the confederation.	confederation peer-as <i>as-number-list</i>	Not configured by default.

Configuring confederation compatibility

If some other routers in the confederation do not comply with RFC 3065, enable confederation compatibility to allow the router to work with those routers.

To configure confederation compatibility:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A
3. Enable compatibility with routers not compliant with RFC 3065 in the confederation.	confederation nonstandard	Optional. Not enabled by default.

Configuring BGP GR

Perform the following configuration on the GR Restarter and GR Helper.

Follow these guidelines when you configure BGP GR:

- A device can act as a GR Restarter and GR Helper at the same time.
- The maximum time allowed for the peer (the GR restarter) to reestablish a BGP session must be less than the Holdtime carried in the Open message.
- The End-Of-RIB (End of Routing-Information-Base) indicates the end of route updates.

To configure BGP GR:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable BGP and enter its view.	bgp as-number	N/A
3. Enable GR Capability for BGP.	graceful-restart	Disabled by default.
4. Configure the maximum time allowed for the peer to reestablish a BGP session.	graceful-restart timer restart timer	Optional. 150 seconds by default.
5. Configure the maximum time to wait for the End-of-RIB marker.	graceful-restart timer wait-for-rib timer	Optional. 180 seconds by default.

Enabling trap

After trap is enabled for BGP, BGP generates Level-4 traps to report important events. The generated traps are sent to the information center of the device. The trap output rules (whether to output the traps and the output direction) are determined according to the information center configuration. (For information center configuration, see *Network Management and Monitoring Configuration Guide*.)

To enable trap:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable trap for BGP.	snmp-agent trap enable bgp	Optional. Enabled by default.

Enabling logging of peer state changes

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enable the logging of peer state changes globally.	log-peer-change	Optional. Enabled by default.
4. Enable the logging of peer state changes for a peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> } log-change	Optional. Enabled by default.

Configuring BFD for BGP

BGP maintains neighbor relationships based on the keepalive timer and holdtime timer, which are set in seconds. BGP defines that the holdtime interval must be at least three times the keepalive interval. This slows down link failure detection; once a failure occurs on a high-speed link, a large quantity of packets will be dropped. BFD is introduced to solve this problem. It detects links between neighbors quickly to reduce convergence time upon link failures.

After a link failure occurs, BFD may detect the failure before the system performs GR. As a result, GR will fail. If GR capability is enabled for BGP, use BFD with caution. If GR and BFD are both enabled, do not disable BFD during a GR process; otherwise, GR may fail. For BFD configuration, see *High Availability Configuration Guide*.

⚠ IMPORTANT:

Before configuring BFD for BGP, you must enable BGP.

To enable BFD for a BGP peer:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enable BFD for the specified BGP peer.	peer <i>ip-address</i> bfd	Not enabled for any BGP peer by default.

Displaying and maintaining BGP

Displaying BGP

Task	Command	Remarks
Display peer group information.	display bgp group [<i>group-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display advertised BGP routing information.	display bgp network [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display AS path information.	display bgp paths [<i>as-regular-expression</i> { begin exclude include } <i>regular-expression</i>]	Available in any view
Display BGP peer or peer group information.	display bgp peer [<i>ip-address</i> { log-info verbose } <i>group-name</i> log-info verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the prefix information in the ORF packet from the specified BGP peer.	display bgp peer <i>ip-address</i> received ip-prefix [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display BGP routing information.	display bgp routing-table [<i>ip-address</i> [{ <i>mask</i> <i>mask-length</i> } [longer-prefixes]]] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display routing information matching the AS path ACL.	display bgp routing-table as-path-acl <i>as-path-acl-number</i> [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display BGP CIDR routing information.	display bgp routing-table cidr [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display BGP routing information matching the specified BGP community.	display bgp routing-table community [<i>aa:nn</i> <1-13>] [no-advertise no-export no-export-subconfed] * [whole-match] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display routing information matching a BGP community list.	display bgp routing-table community-list { { <i>basic-community-list-number</i> <i>comm-list-name</i> } [whole-match] <i>adv-community-list-number</i> } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display BGP dampened routing information.	display bgp routing-table dampened [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display BGP dampening parameter information.	display bgp routing-table dampening parameter [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display BGP routing information originating from different ASs.	display bgp routing-table different-origin-as [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display BGP routing flap statistics.	display bgp routing-table flap-info [<i>regular-expression as-regular-expression</i> [as-path-acl <i>as-path-acl-number</i> <i>ip-address</i> [{ <i>mask</i> <i>mask-length</i> } [longer-match]]]] [{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display labeled BGP routing information.	display bgp routing-table label [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display routing information to or from a peer.	display bgp routing-table peer <i>ip-address</i> { advertised-routes received-routes } [<i>network-address</i> [<i>mask</i> <i>mask-length</i>]] statistic] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display routing information matching a regular expression.	display bgp routing-table regular-expression <i>as-regular-expression</i>	Available in any view

Task	Command	Remarks
Display BGP routing statistics.	display bgp routing-table statistic [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the global router ID.	display router id [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Resetting BGP connections

Task	Command	Remarks
Reset all BGP connections.	reset bgp all	Available in user view
Reset the BGP connections to an AS.	reset bgp as-number	Available in user view
Reset the BGP connection to a peer.	reset bgp ip-address [flap-info]	Available in user view
Reset all EBGP connections.	reset bgp external	Available in user view
Reset the BGP connections to a peer group.	reset bgp group <i>group-name</i>	Available in user view
Reset all IBGP connections.	reset bgp internal	Available in user view
Reset all IPv4 unicast BGP connections.	reset bgp ipv4 all	Available in user view

Clearing BGP information

Task	Command	Remarks
Clear dampened BGP routing information and release suppressed routes.	reset bgp dampening [<i>ip-address</i> [<i>mask</i> <i>mask-length</i>]]	Available in user view
Clear route flap information.	reset bgp flap-info [<i>ip-address</i> [<i>mask-length</i> <i>mask</i>] as-path-acl <i>as-path-acl-number</i> regex <i>as-path-regular-expression</i>]	Available in user view

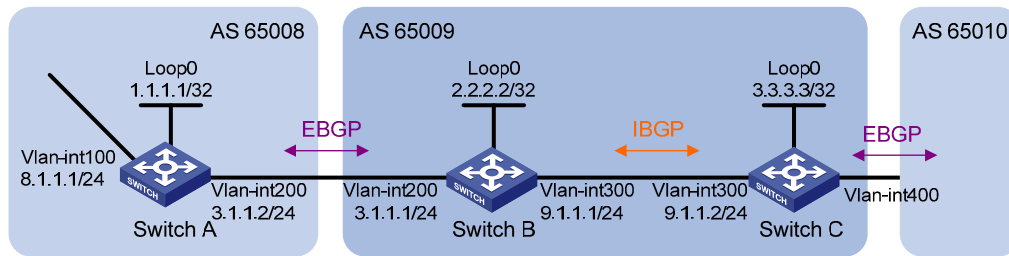
BGP configuration examples

BGP basic configuration

Network requirements

In [Figure 90](#), run EBGP between Switch A and Switch B and IBGP between Switch B and Switch C so that Switch C can access the network 8.1.1.0/24 connected to Router A.

Figure 90 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure IBGP:
 - To prevent route flapping caused by port state changes, this example uses loopback interfaces to establish IBGP connections.
 - Because loopback interfaces are virtual interfaces, you need to use the **peer connect-interface** command to specify the loopback interface as the source interface for establishing BGP connections.
 - Enable OSPF in AS 65009 to make sure that Switch B can communicate with Switch C through loopback interfaces.

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] bgp 65009
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] peer 3.3.3.3 as-number 65009
[SwitchB-bgp] peer 3.3.3.3 connect-interface loopback 0
[SwitchB-bgp] quit
[SwitchB] ospf 1
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 2.2.2.2 0.0.0.0
[SwitchB-ospf-1-area-0.0.0.0] network 9.1.1.1 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] bgp 65009
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] peer 2.2.2.2 as-number 65009
[SwitchC-bgp] peer 2.2.2.2 connect-interface loopback 0
[SwitchC-bgp] quit
[SwitchC] ospf 1
[SwitchC-ospf-1] area 0
[SwitchC-ospf-1-area-0.0.0.0] network 3.3.3.3 0.0.0.0
[SwitchC-ospf-1-area-0.0.0.0] network 9.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] quit
[SwitchC-ospf-1] quit
[SwitchC] display bgp peer
```



```

BGP local router ID : 3.3.3.3
Local AS number : 65009
Total number of peers : 1                Peers in established state : 1

Peer                AS  MsgRcvd  MsgSent  OutQ  PrefRcv  Up/Down  State
2.2.2.2            65009    2        2    0      0 00:00:13 Established

```

The output shows that Switch C has established an IBGP peer relationship with Switch B.

3. Configure EBGP:

- The EBGP peers, Switch A and Switch B (usually belong to different carriers), are located in different ASs. Typically, their loopback interfaces are not reachable to each other, so directly connected interfaces are used for establishing BGP sessions.
- To enable Switch C to access the network 8.1.1.0/24 connected directly to Switch A, inject network 8.1.1.0/24 to the BGP routing table of Switch A.

Configure Switch A.

```

<SwitchA> system-view
[SwitchA] bgp 65008
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] peer 3.1.1.1 as-number 65009
[SwitchA-bgp] network 8.1.1.1 24
[SwitchA-bgp] quit

```

Configure Switch B.

```

[SwitchB] bgp 65009
[SwitchB-bgp] peer 3.1.1.2 as-number 65008
[SwitchB-bgp] quit

```

Display BGP peer information on Switch B.

```

[SwitchB] display bgp peer

```

```

BGP local router ID : 2.2.2.2
Local AS number : 65009
Total number of peers : 2                Peers in established state : 2

Peer                AS  MsgRcvd  MsgSent  OutQ  PrefRcv  Up/Down  State
3.3.3.3            65009    4        4    0      0 00:02:49 Established
3.1.1.2            65008    2        2    0      0 00:00:05 Established

```

The output shows that Switch B has established an IBGP peer relationship with Switch C and an EBGP peer relationship with Switch A.

Display the BGP routing table on Switch A.

```

[SwitchA] display bgp routing-table

```

```

Total Number of Routes: 1

```

```

BGP Local router ID is 1.1.1.1

```

```

Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
              h - history, i - internal, s - suppressed, S - Stale

```

Origin : i - IGP, e - EGP, ? - incomplete

Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*> 8.1.1.0/24	0.0.0.0	0		0	i

Display the BGP routing table on Switch B.

[SwitchB] display bgp routing-table

Total Number of Routes: 1

BGP Local router ID is 2.2.2.2

Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
h - history, i - internal, s - suppressed, S - Stale
Origin : i - IGP, e - EGP, ? - incomplete

Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*> 8.1.1.0/24	3.1.1.2	0		0	65008i

Display the BGP routing table on Switch C.

[SwitchC] display bgp routing-table

Total Number of Routes: 1

BGP Local router ID is 3.3.3.3

Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
h - history, i - internal, s - suppressed, S - Stale
Origin : i - IGP, e - EGP, ? - incomplete

Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
i 8.1.1.0/24	3.1.1.2	0	100	0	65008i

The outputs show that Switch A has learned no route to AS65009, and Switch C has learned network 8.1.1.0 but the next hop 3.1.1.2 is unreachable, so the route is invalid.

4. Redistribute direct routes:

Configure BGP to redistribute direct routes on Switch B, so that Switch A can obtain the route to 9.1.1.0/24 and Switch C can obtain the route to 3.1.1.0/24.

Configure Switch B.

[SwitchB] bgp 65009

[SwitchB-bgp] import-route direct

Display the BGP routing table on Switch A.

[SwitchA] display bgp routing-table

Total Number of Routes: 4

BGP Local router ID is 1.1.1.1

Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
h - history, i - internal, s - suppressed, S - Stale

Origin : i - IGP, e - EGP, ? - incomplete

	Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>	2.2.2.2/32	3.1.1.1	0		0	65009?
*	3.1.1.0/24	3.1.1.1	0		0	65009?
*>	8.1.1.0/24	0.0.0.0	0		0	i
*>	9.1.1.0/24	3.1.1.1	0		0	65009?

Two routes 2.2.2.2/32 and 9.1.1.0/24 have been added in Switch A's routing table.

Display the BGP routing table on Switch C.

```
[SwitchC] display bgp routing-table
```

Total Number of Routes: 4

BGP Local router ID is 3.3.3.3

Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
h - history, i - internal, s - suppressed, S - Stale
Origin : i - IGP, e - EGP, ? - incomplete

	Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
i	2.2.2.2/32	2.2.2.2	0	100	0	?
*>i	3.1.1.0/24	2.2.2.2	0	100	0	?
*>i	8.1.1.0/24	3.1.1.2	0	100	0	65008i
* i	9.1.1.0/24	2.2.2.2	0	100	0	?

The output shows that the route 8.1.1.0 becomes valid with the next hop as Switch A.

5. Verify the configuration:

Ping 8.1.1.1 on Switch C.

```
[SwitchC] ping 8.1.1.1
```

```
PING 8.1.1.1: 56 data bytes, press CTRL_C to break
Reply from 8.1.1.1: bytes=56 Sequence=1 ttl=254 time=2 ms
Reply from 8.1.1.1: bytes=56 Sequence=2 ttl=254 time=2 ms
Reply from 8.1.1.1: bytes=56 Sequence=3 ttl=254 time=2 ms
Reply from 8.1.1.1: bytes=56 Sequence=4 ttl=254 time=2 ms
Reply from 8.1.1.1: bytes=56 Sequence=5 ttl=254 time=2 ms
--- 8.1.1.1 ping statistics ---
5 packet(s) transmitted
5 packet(s) received
0.00% packet loss
round-trip min/avg/max = 2/2/2 ms
```

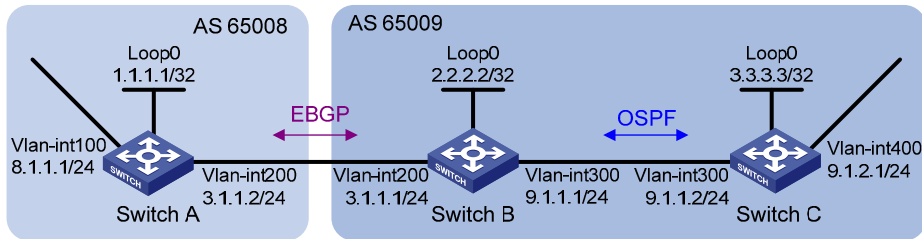
BGP and IGP synchronization configuration

Network requirements

As shown in [Figure 91](#), all devices of company A belong to AS 65008, and all devices of company B belong to AS 65009. AS 65008 and AS 65009 are connected through Switch A and Switch B. Switch

A must be able to access network 9.1.2.0/24 in AS 65009, and Switch C must access network 8.1.1.0/24 in AS 65008.

Figure 91 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure OSPF:

Enable OSPF in AS 65009, so that Switch B can obtain the route to 9.1.2.0/24.

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] ospf 1
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 2.2.2.2 0.0.0.0
[SwitchB-ospf-1-area-0.0.0.0] network 9.1.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] ospf 1
[SwitchC-ospf-1] import-route direct
[SwitchC-ospf-1] area 0
[SwitchC-ospf-1-area-0.0.0.0] network 9.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] quit
[SwitchC-ospf-1] quit
```

3. Configure the EBGp connection:

Configure the EBGp connection and inject network 8.1.1.0/24 to the BGP routing table of Switch A, so that Switch B can obtain the route to 8.1.1.0/24.

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] bgp 65008
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] peer 3.1.1.1 as-number 65009
[SwitchA-bgp] network 8.1.1.0 24
[SwitchA-bgp] quit
```

Configure Switch B.

```
[SwitchB] bgp 65009
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] peer 3.1.1.2 as-number 65008
```

4. Configure BGP and IGP synchronization:

- Configure BGP to redistribute routes from OSPF on Switch B, so that Switch A can obtain the route to 9.1.2.0/24.
- Configure OSPF to redistribute routes from BGP on Switch B, so that Switch C can obtain the route to 8.1.1.0/24.

Configure BGP to redistribute routes from OSPF on Switch B.

```
[SwitchB-bgp] import-route ospf 1
[SwitchB-bgp] quit
[SwitchB] ospf 1
[SwitchB-ospf-1] import-route bgp
[SwitchB-ospf-1] quit
```

Display the BGP routing table on Switch A.

```
[SwitchA] display bgp routing-table
```

Total Number of Routes: 3

BGP Local router ID is 1.1.1.1

Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,

h - history, i - internal, s - suppressed, S - Stale

Origin : i - IGP, e - EGP, ? - incomplete

	Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>	3.3.3.3/32	3.1.1.1	1		0	65009?
*>	8.1.1.0/24	0.0.0.0	0		0	i
*>	9.1.2.0/24	3.1.1.1	1		0	65009?

Display the routing table on Switch C.

```
[SwitchC] display ip routing-table
```

Routing Tables: Public

Destinations : 9 Routes : 9

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
2.2.2.2/32	OSPF	10	1	9.1.1.1	Vlan300
3.3.3.3/32	Direct	0	0	127.0.0.1	InLoop0
8.1.1.0/24	O_ASE	150	1	9.1.1.1	Vlan300
9.1.1.0/24	Direct	0	0	9.1.1.2	Vlan300
9.1.1.2/32	Direct	0	0	127.0.0.1	InLoop0
9.1.2.0/24	Direct	0	0	9.1.2.1	Vlan400
9.1.2.1/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0

5. Verify the configuration:

Use ping for verification.

```
[SwitchA] ping -a 8.1.1.1 9.1.2.1
```

PING 9.1.2.1: 56 data bytes, press CTRL_C to break

Reply from 9.1.2.1: bytes=56 Sequence=1 ttl=254 time=15 ms

```

Reply from 9.1.2.1: bytes=56 Sequence=2 ttl=254 time=31 ms
Reply from 9.1.2.1: bytes=56 Sequence=3 ttl=254 time=47 ms
Reply from 9.1.2.1: bytes=56 Sequence=4 ttl=254 time=46 ms
Reply from 9.1.2.1: bytes=56 Sequence=5 ttl=254 time=47 ms

--- 9.1.2.1 ping statistics ---
 5 packet(s) transmitted
 5 packet(s) received
 0.00% packet loss
 round-trip min/avg/max = 15/37/47 ms
[SwitchC] ping -a 9.1.2.1 8.1.1.1
PING 8.1.1.1: 56 data bytes, press CTRL_C to break
Reply from 8.1.1.1: bytes=56 Sequence=1 ttl=254 time=2 ms
Reply from 8.1.1.1: bytes=56 Sequence=2 ttl=254 time=2 ms
Reply from 8.1.1.1: bytes=56 Sequence=3 ttl=254 time=2 ms
Reply from 8.1.1.1: bytes=56 Sequence=4 ttl=254 time=2 ms
Reply from 8.1.1.1: bytes=56 Sequence=5 ttl=254 time=2 ms

--- 8.1.1.1 ping statistics ---
 5 packet(s) transmitted
 5 packet(s) received
 0.00% packet loss
 round-trip min/avg/max = 2/2/2 ms

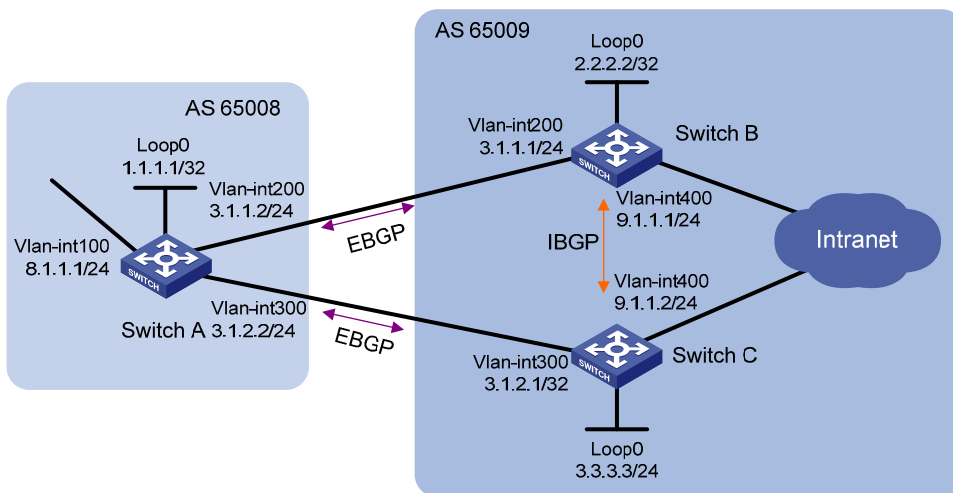
```

BGP load balancing configuration

Network requirements

As shown in [Figure 92](#), all the switches run BGP. Switch A resides in AS 65008, Switch B and Switch C in AS 65009. Between Switch A and Switch B, Switch A and Switch C are EBGP connections, and between Switch B and Switch C is an IBGP connection. Two routes are configured on Switch A for load balancing.

Figure 92 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure BGP connections:
 - On Switch A, establish EBGP connections with Switch B and Switch C; configure BGP to advertise network 8.1.1.0/24 to Switch B and Switch C, so that Switch B and Switch C can access the internal network connected to Switch A.
 - On Switch B, establish an EBGP connection with Switch A and an IBGP connection with Switch C; configure BGP to advertise network 9.1.1.0/24 to Switch A, so that Switch A can access the intranet through Switch B; configure a static route to interface loopback 0 on Switch C (or use a routing protocol like OSPF) to establish the IBGP connection.
 - On Switch C, establish an EBGP connection with Switch A and an IBGP connection with Switch B; configure BGP to advertise network 9.1.1.0/24 to Switch A, so that Switch A can access the intranet through Switch C; configure a static route to interface loopback 0 on Switch B (or use another protocol like OSPF) to establish the IBGP connection.

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] bgp 65008
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] peer 3.1.1.1 as-number 65009
[SwitchA-bgp] peer 3.1.2.1 as-number 65009
[SwitchA-bgp] network 8.1.1.1 24
[SwitchA-bgp] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] bgp 65009
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] peer 3.1.1.2 as-number 65008
[SwitchB-bgp] peer 3.3.3.3 as-number 65009
[SwitchB-bgp] peer 3.3.3.3 connect-interface loopback 0
[SwitchB-bgp] network 9.1.1.0 255.255.255.0
[SwitchB-bgp] quit
[SwitchB] ip route-static 3.3.3.3 32 9.1.1.2
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] bgp 65009
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] peer 3.1.2.2 as-number 65008
[SwitchC-bgp] peer 2.2.2.2 as-number 65009
[SwitchC-bgp] peer 2.2.2.2 connect-interface loopback 0
[SwitchC-bgp] network 9.1.1.0 255.255.255.0
[SwitchC-bgp] quit
[SwitchC] ip route-static 2.2.2.2 32 9.1.1.1
```

Display the BGP routing table on Switch A.

```
[SwitchA] display bgp routing-table
```

```
Total Number of Routes: 3
```

BGP Local router ID is 1.1.1.1

Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
h - history, i - internal, s - suppressed, S - Stale
Origin : i - IGP, e - EGP, ? - incomplete

	Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>	8.1.1.0/24	0.0.0.0	0		0	i
*>	9.1.1.0/24	3.1.1.1	0		0	65009i
*		3.1.2.1	0		0	65009i

- The output shows two valid routes to destination 9.1.1.0/24: the route with next hop 3.1.1.1 is marked with a greater-than sign (>), indicating it is the best route (because the ID of Switch B is smaller); the route with next hop 3.1.2.1 is marked with only an asterisk (*), indicating it is a valid route, but not the best.
- By using the **display ip routing-table** command, you can find only one route to 9.1.1.0/24 with next hop 3.1.1.1 and outbound interface VLAN-interface 200.

3. Configure loading balancing:

Since Switch A has two routes to reach AS 65009, configuring load balancing over the two BGP routes on Switch A can improve link utilization.

Configure Switch A.

```
[SwitchA] bgp 65008
[SwitchA-bgp] balance 2
[SwitchA-bgp] quit
```

4. Verify the configuration:

Display the BGP routing table on Switch A.

```
[SwitchA] display bgp routing-table
```

Total Number of Routes: 3

BGP Local router ID is 1.1.1.1

Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
h - history, i - internal, s - suppressed, S - Stale
Origin : i - IGP, e - EGP, ? - incomplete

	Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>	8.1.1.0/24	0.0.0.0	0		0	i
*>	9.1.1.0/24	3.1.1.1	0		0	65009i
*>		3.1.2.1	0		0	65009i

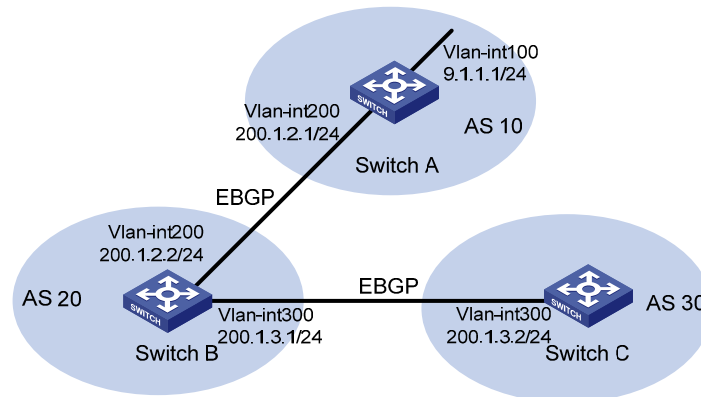
- The route 9.1.1.0/24 has two next hops 3.1.1.1 and 3.1.2.1, both of which are marked with a greater-than sign (>), indicating they are the best routes.
- By using the **display ip routing-table** command, you can find two routes to 9.1.1.0/24: one with next hop 3.1.1.1 and outbound interface VLAN-interface 200, the other with next hop 3.1.2.1 and outbound interface VLAN-interface 300.

BGP community configuration

Network requirements

As shown in Figure 93, Switch B establishes EBGP connections with Switch A and C. Configure No_Export community attribute on Switch A to make routes from AS 10 not advertised by AS 20 to any other AS.

Figure 93 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure EBGP:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] bgp 10
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] peer 200.1.2.2 as-number 20
[SwitchA-bgp] network 9.1.1.0 255.255.255.0
[SwitchA-bgp] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] bgp 20
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] peer 200.1.2.1 as-number 10
[SwitchB-bgp] peer 200.1.3.2 as-number 30
[SwitchB-bgp] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] bgp 30
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] peer 200.1.3.1 as-number 20
[SwitchC-bgp] quit
```

Display the BGP routing table on Switch B.

```
[SwitchB] display bgp routing-table 9.1.1.0
```

```
BGP local router ID : 2.2.2.2
```

```
Local AS number : 20
Paths: 1 available, 1 best
```

BGP routing table entry information of 9.1.1.0/24:

```
From : 200.1.2.1 (1.1.1.1)
Original nexthop: 200.1.2.1
AS-path : 10
Origin : igp
Attribute value : MED 0, pref-val 0, pre 255
State : valid, external, best,
Advertised to such 1 peers:
    200.1.3.2
```

Switch B advertised routes to Switch C in AS30.

Display the routing table on Switch C.

```
[SwitchC] display bgp routing-table
```

```
Total Number of Routes: 1
```

```
BGP Local router ID is 3.3.3.3
```

```
Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
              h - history, i - internal, s - suppressed, S - Stale
              Origin : i - IGP, e - EGP, ? - incomplete
```

Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*> 9.1.1.0/24	200.1.3.1	0		0	20 10i

Switch C has learned route 9.1.1.0/24 from Switch B.

3. Configure BGP community:

Configure a routing policy.

```
[SwitchA] route-policy comm_policy permit node 0
[SwitchA-route-policy] apply community no-export
[SwitchA-route-policy] quit
```

Apply the routing policy.

```
[SwitchA] bgp 10
[SwitchA-bgp] peer 200.1.2.2 route-policy comm_policy export
[SwitchA-bgp] peer 200.1.2.2 advertise-community
```

Display the routing table on Switch B.

```
[SwitchB] display bgp routing-table 9.1.1.0
BGP local router ID : 2.2.2.2
Local AS number : 20
Paths: 1 available, 1 best
```

BGP routing table entry information of 9.1.1.0/24:

```
From : 200.1.2.1 (1.1.1.1)
Original nexthop: 200.1.2.1
Community : No-Export
AS-path : 10
Origin : igp
```

Attribute value : MED 0, pref-val 0, pre 255
State : valid, external, best,
Not advertised to any peers yet

The route 9.1.1.0/24 is not available in the routing table of Switch C.

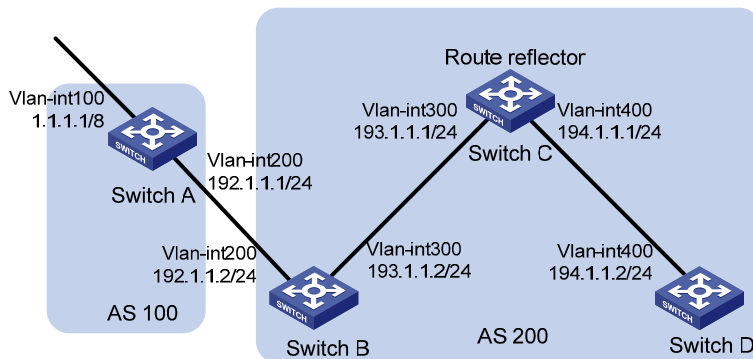
BGP route reflector configuration

Network requirements

In Figure 94, all switches run BGP.

- Between Switch A and Switch B is an EBGP connection; between Switch C and Switch B, and between Switch C and Switch D are IBGP connections.
- Switch C is a route reflector with clients Switch B and D.
- Switch D can learn route 1.0.0.0/8 from Switch C.

Figure 94 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure BGP connections:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] bgp 100
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] peer 192.1.1.2 as-number 200
```

Inject network 1.0.0.0/8 to the BGP routing table.

```
[SwitchA-bgp] network 1.0.0.0
[SwitchA-bgp] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] bgp 200
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] peer 192.1.1.1 as-number 100
[SwitchB-bgp] peer 193.1.1.1 as-number 200
[SwitchB-bgp] peer 193.1.1.1 next-hop-local
[SwitchB-bgp] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] bgp 200
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] peer 193.1.1.2 as-number 200
[SwitchC-bgp] peer 194.1.1.2 as-number 200
[SwitchC-bgp] quit
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] bgp 200
[SwitchD-bgp] router-id 4.4.4.4
[SwitchD-bgp] peer 194.1.1.1 as-number 200
[SwitchD-bgp] quit
```

3. Configure the route reflector:

Configure Switch C.

```
[SwitchC] bgp 200
[SwitchC-bgp] peer 193.1.1.2 reflect-client
[SwitchC-bgp] peer 194.1.1.2 reflect-client
[SwitchC-bgp] quit
```

4. Verify the configuration:

Display the BGP routing table on Switch B.

```
[SwitchB] display bgp routing-table
```

Total Number of Routes: 1

BGP Local router ID is 200.1.2.2

Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
h - history, i - internal, s - suppressed, S - Stale
Origin : i - IGP, e - EGP, ? - incomplete

Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*> 1.0.0.0	192.1.1.1	0		0	100i

Display the BGP routing table on Switch D.

```
[SwitchD] display bgp routing-table
```

Total Number of Routes: 1

BGP Local router ID is 200.1.2.1

Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
h - history, i - internal, s - suppressed, S - Stale
Origin : i - IGP, e - EGP, ? - incomplete

Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
i 1.0.0.0	193.1.1.2	0	100	0	100i

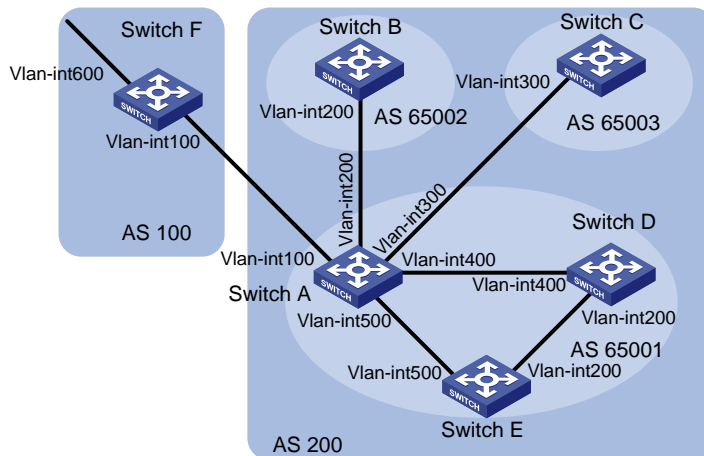
Switch D has learned route 1.0.0.0/8 from Switch C.

BGP confederation configuration

Network requirements

In Figure 95, to reduce IBGP connections in AS 200, split it into three sub-ASs: AS65001, AS65002, and AS65003. Switches in AS65001 are fully meshed.

Figure 95 Network diagram



Device	Interface	IP address	Device	Interface	IP address
Switch A	Vlan-int100	200.1.1.1/24	Switch D	Vlan-int200	10.1.5.1/24
	Vlan-int200	10.1.1.1/24		Vlan-int400	10.1.3.2/24
	Vlan-int300	10.1.2.1/24	Switch E	Vlan-int200	10.1.5.2/24
	Vlan-int400	10.1.3.1/24		Vlan-int500	10.1.4.2/24
	Vlan-int500	10.1.4.1/24		Switch F	Vlan-int100
Switch B	Vlan-int200	10.1.1.2/24	Vlan-int600		9.1.1.1/24
Switch C	Vlan-int300	10.1.2.2/24			

Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure BGP confederation:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] bgp 65001
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] confederation id 200
[SwitchA-bgp] confederation peer-as 65002 65003
[SwitchA-bgp] peer 10.1.1.2 as-number 65002
[SwitchA-bgp] peer 10.1.1.2 next-hop-local
[SwitchA-bgp] peer 10.1.2.2 as-number 65003
[SwitchA-bgp] peer 10.1.2.2 next-hop-local
[SwitchA-bgp] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] bgp 65002
[SwitchB-bgp] router-id 2.2.2.2
```

```
[SwitchB-bgp] confederation id 200
[SwitchB-bgp] confederation peer-as 65001 65003
[SwitchB-bgp] peer 10.1.1.1 as-number 65001
[SwitchB-bgp] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] bgp 65003
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] confederation id 200
[SwitchC-bgp] confederation peer-as 65001 65002
[SwitchC-bgp] peer 10.1.2.1 as-number 65001
[SwitchC-bgp] quit
```

3. Configure IBGP connections in AS65001:

Configure Switch A.

```
[SwitchA] bgp 65001
[SwitchA-bgp] peer 10.1.3.2 as-number 65001
[SwitchA-bgp] peer 10.1.3.2 next-hop-local
[SwitchA-bgp] peer 10.1.4.2 as-number 65001
[SwitchA-bgp] peer 10.1.4.2 next-hop-local
[SwitchA-bgp] quit
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] bgp 65001
[SwitchD-bgp] router-id 4.4.4.4
[SwitchD-bgp] confederation id 200
[SwitchD-bgp] peer 10.1.3.1 as-number 65001
[SwitchD-bgp] peer 10.1.5.2 as-number 65001
[SwitchD-bgp] quit
```

Configure Switch E.

```
<SwitchE> system-view
[SwitchE] bgp 65001
[SwitchE-bgp] router-id 5.5.5.5
[SwitchE-bgp] confederation id 200
[SwitchE-bgp] peer 10.1.4.1 as-number 65001
[SwitchE-bgp] peer 10.1.5.1 as-number 65001
[SwitchE-bgp] quit
```

4. Configure the EBGP connection between AS100 and AS200:

Configure Switch A.

```
[SwitchA] bgp 65001
[SwitchA-bgp] peer 200.1.1.2 as-number 100
[SwitchA-bgp] quit
```

Configure Switch F.

```
<SwitchF> system-view
[SwitchF] bgp 100
[SwitchF-bgp] router-id 6.6.6.6
[SwitchF-bgp] peer 200.1.1.1 as-number 200
[SwitchF-bgp] network 9.1.1.0 255.255.255.0
```

```
[SwitchF-bgp] quit
```

5. Verify the configuration:

```
# Display the routing table on Switch B.
```

```
[SwitchB] display bgp routing-table
```

```
Total Number of Routes: 1
```

```
BGP Local router ID is 2.2.2.2
```

```
Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,  
              h - history, i - internal, s - suppressed, S - Stale  
              Origin : i - IGP, e - EGP, ? - incomplete
```

Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>i 9.1.1.0/24	10.1.1.1	0	100	0	(65001) 100i

```
[SwitchB] display bgp routing-table 9.1.1.0
```

```
BGP local router ID : 2.2.2.2
```

```
Local AS number : 65002
```

```
Paths: 1 available, 1 best
```

```
BGP routing table entry information of 9.1.1.0/24:
```

```
From : 10.1.1.1 (1.1.1.1)  
Relay Nexthop : 0.0.0.0  
Original nexthop: 10.1.1.1  
AS-path : (65001) 100  
Origin : igp  
Attribute value : MED 0, localpref 100, pref-val 0, pre 255  
State : valid, external-confed, best,  
Not advertised to any peers yet
```

```
# Display the BGP routing table on Switch D.
```

```
[SwitchD] display bgp routing-table
```

```
Total Number of Routes: 1
```

```
BGP Local router ID is 4.4.4.4
```

```
Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,  
              h - history, i - internal, s - suppressed, S - Stale  
              Origin : i - IGP, e - EGP, ? - incomplete
```

Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>i 9.1.1.0/24	10.1.3.1	0	100	0	100i

```
[SwitchD] display bgp routing-table 9.1.1.0
```

```
BGP local router ID : 4.4.4.4
```

```
Local AS number : 65001
```

```
Paths: 1 available, 1 best
```

```
BGP routing table entry information of 9.1.1.0/24:
```

```

From          : 10.1.3.1 (1.1.1.1)
Relay Nexthop : 0.0.0.0
Original nexthop: 10.1.3.1
AS-path       : 100
Origin        : igp
Attribute value : MED 0, localpref 100, pref-val 0, pre 255
State         : valid, internal, best,
Not advertised to any peers yet

```

The output shows the following:

- Switch F can send route information to Switch B and Switch C through the confederation by establishing only an EBGP connection with Switch A.
- Switch B and Switch D are in the same confederation, but belong to different sub ASs. They obtain external route information from Switch A and generate the same BGP route entries; it seems like they reside in the same AS although they have no direct connection in between.

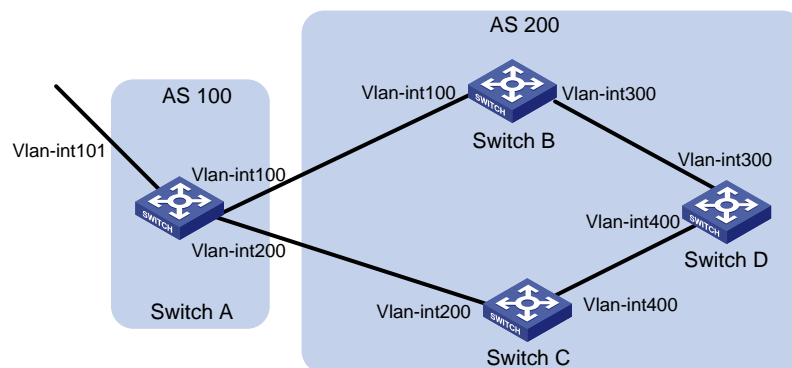
BGP path selection configuration

Network requirements

In [Figure 96](#), all switches run BGP. Between Switch A and Switch B, and between Switch A and Switch C are EBGP connections. Between Switch B and Switch D, and between Switch D and Switch C are IBGP connections. OSPF is the IGP protocol in AS 200.

Configure routing policies, making Switch D use the route 1.0.0.0/8 from Switch C as the optimal.

Figure 96 Network diagram



Device	Interface	IP address	Device	Interface	IP address
Switch A	Vlan-int101	1.0.0.0/8	Switch D	Vlan-int400	195.1.1.1/24
	Vlan-int100	192.1.1.1/24		Vlan-int300	194.1.1.1/24
	Vlan-int200	193.1.1.1/24	Switch C	Vlan-int400	195.1.1.2/24
Switch B	Vlan-int100	192.1.1.2/24		Vlan-int200	193.1.1.2/24
	Vlan-int300	194.1.1.2/24			

Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure OSPF on Switch B, C, and D:

Configure Switch B.

```
<SwitchB> system-view
```



```
[SwitchB] ospf
[SwitchB-ospf] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 192.1.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] network 194.1.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] ospf
[SwitchC-ospf] area 0
[SwitchC-ospf-1-area-0.0.0.0] network 193.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] network 195.1.1.0 0.0.0.255
[SwitchC-ospf-1-area-0.0.0.0] quit
[SwitchC-ospf-1] quit
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] ospf
[SwitchD-ospf] area 0
[SwitchD-ospf-1-area-0.0.0.0] network 194.1.1.0 0.0.0.255
[SwitchD-ospf-1-area-0.0.0.0] network 195.1.1.0 0.0.0.255
[SwitchD-ospf-1-area-0.0.0.0] quit
[SwitchD-ospf-1] quit
```

3. Configure BGP connections:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] bgp 100
[SwitchA-bgp] peer 192.1.1.2 as-number 200
[SwitchA-bgp] peer 193.1.1.2 as-number 200
```

Inject network 1.0.0.0/8 to the BGP routing table on Switch A.

```
[SwitchA-bgp] network 1.0.0.0 8
[SwitchA-bgp] quit
```

Configure Switch B.

```
[SwitchB] bgp 200
[SwitchB-bgp] peer 192.1.1.1 as-number 100
[SwitchB-bgp] peer 194.1.1.1 as-number 200
[SwitchB-bgp] quit
```

Configure Switch C.

```
[SwitchC] bgp 200
[SwitchC-bgp] peer 193.1.1.1 as-number 100
[SwitchC-bgp] peer 195.1.1.1 as-number 200
[SwitchC-bgp] quit
```

Configure Switch D.

```
[SwitchD] bgp 200
[SwitchD-bgp] peer 194.1.1.2 as-number 200
[SwitchD-bgp] peer 195.1.1.2 as-number 200
[SwitchD-bgp] quit
```

4. Configure attributes for route 1.0.0.0/8, making Switch D give priority to the route learned from Switch C:

o Method I

Configure a higher MED value for the route 1.0.0.0/8 advertised from Switch A to peer 192.1.1.2.

Define an ACL numbered 2000 to permit route 1.0.0.0/8.

```
[SwitchA] acl number 2000
[SwitchA-acl-basic-2000] rule permit source 1.0.0.0 0.255.255.255
[SwitchA-acl-basic-2000] quit
```

Define two routing policies, **apply_med_50**, which sets the MED for route 1.0.0.0/8 to 50, and **apply_med_100**, which sets the MED for route 1.0.0.0/8 to 100.

```
[SwitchA] route-policy apply_med_50 permit node 10
[SwitchA-route-policy] if-match acl 2000
[SwitchA-route-policy] apply cost 50
[SwitchA-route-policy] quit
[SwitchA] route-policy apply_med_100 permit node 10
[SwitchA-route-policy] if-match acl 2000
[SwitchA-route-policy] apply cost 100
[SwitchA-route-policy] quit
```

Apply routing policy **apply_med_50** to the route advertised to peer 193.1.1.2 (Switch C), and **apply_med_100** to the route advertised to peer 192.1.1.2 (Switch B).

```
[SwitchA] bgp 100
[SwitchA-bgp] peer 193.1.1.2 route-policy apply_med_50 export
[SwitchA-bgp] peer 192.1.1.2 route-policy apply_med_100 export
[SwitchA-bgp] quit
```

Display the BGP routing table on Switch D.

```
[SwitchD] display bgp routing-table
```

```
Total Number of Routes: 2
```

```
BGP Local router ID is 194.1.1.1
```

```
Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
```

```
h - history, i - internal, s - suppressed, S - Stale
```

```
Origin : i - IGP, e - EGP, ? - incomplete
```

Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>i 1.0.0.0	193.1.1.1	50	100	0	100i
* i	192.1.1.1	100	100	0	100i

Route 1.0.0.0/8 is the optimal.

o Method II

Configure different local preferences on Switch B and C for route 1.0.0.0/8, making Switch D give priority to the route from Switch C.

Define an ACL numbered 2000 on Router C, permitting route 1.0.0.0/8.

```
[SwitchC] acl number 2000
[SwitchC-acl-basic-2000] rule permit source 1.0.0.0 0.255.255.255
[SwitchC-acl-basic-2000] quit
```

```
# Configure a routing policy named localpref on Switch C, setting the local preference of route 1.0.0.0/8 to 200 (the default is 100).
```

```
[SwitchC] route-policy localpref permit node 10
[SwitchC-route-policy] if-match acl 2000
[SwitchC-route-policy] apply local-preference 200
[SwitchC-route-policy] quit
```

```
# Apply routing policy localpref to routes from peer 193.1.1.1.
```

```
[SwitchC] bgp 200
[SwitchC-bgp] peer 193.1.1.1 route-policy localpref import
[SwitchC-bgp] quit
```

```
# Display the routing table on Switch D.
```

```
[SwitchD] display bgp routing-table
```

```
Total Number of Routes: 2
```

```
BGP Local router ID is 194.1.1.1
```

```
Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
              h - history, i - internal, s - suppressed, S - Stale
              Origin : i - IGP, e - EGP, ? - incomplete
```

Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>i 1.0.0.0	193.1.1.1	0	200	0	100i
* i	192.1.1.1	0	100	0	100i

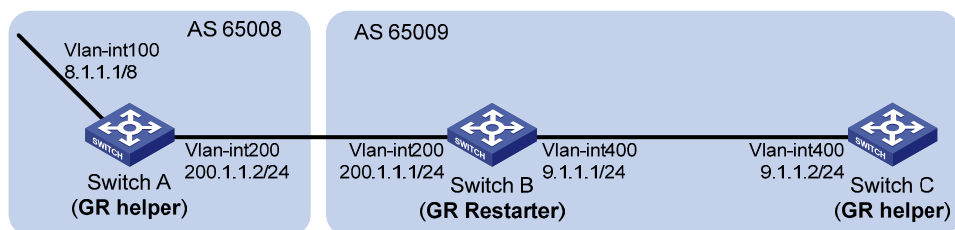
```
Route 1.0.0.0/8 from Switch D to Switch C is the optimal.
```

BGP GR configuration

Network requirements

All switches run BGP in [Figure 97](#). Between Switch A and Switch B is an EBGP connection. Switch B and Switch C are connected over an IBGP connection. Enable GR capability for BGP so that the communication between Switch A and Switch C is not affected when a master/slave switchover occurs on Switch B.

Figure 97 Network diagram



Configuration procedure

1. Configure Switch A:

```
# Configure IP addresses for interfaces. (Details not shown.)
```

```
# Configure the EBGP connection.
```

```
<SwitchA> system-view
```

```
[SwitchA] bgp 65008
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] peer 200.1.1.1 as-number 65009
# Inject network 8.0.0.0/8 to the BGP routing table.
[SwitchA-bgp] network 8.0.0.0
# Enable GR capability for BGP.
[SwitchA-bgp] graceful-restart
```

2. Configure Switch B:

```
# Configure IP addresses for interfaces. (Details not shown.)
# Configure the EBGP connection.
<SwitchB> system-view
[SwitchB] bgp 65009
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] peer 200.1.1.2 as-number 65008
# Configure the IBGP connection.
[SwitchB-bgp] peer 9.1.1.2 as-number 65009
# Inject networks 200.1.1.0/24 and 9.1.1.0/24 to the BGP routing table.
[SwitchB-bgp] network 200.1.1.0 24
[SwitchB-bgp] network 9.1.1.0 24
# Enable GR capability for BGP.
[SwitchB-bgp] graceful-restart
```

3. Configure Switch C:

```
# Configure IP addresses for interfaces. (Details not shown.)
# Configure the IBGP connection.
<SwitchC> system-view
[SwitchC] bgp 65009
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] peer 9.1.1.1 as-number 65009
# Enable GR capability for BGP.
[SwitchC-bgp] graceful-restart
```

4. Verify the configuration:

Ping Switch C on Switch A. Meanwhile, perform a master/slave switchover on Switch B. The ping operation is successful during the whole switchover process.

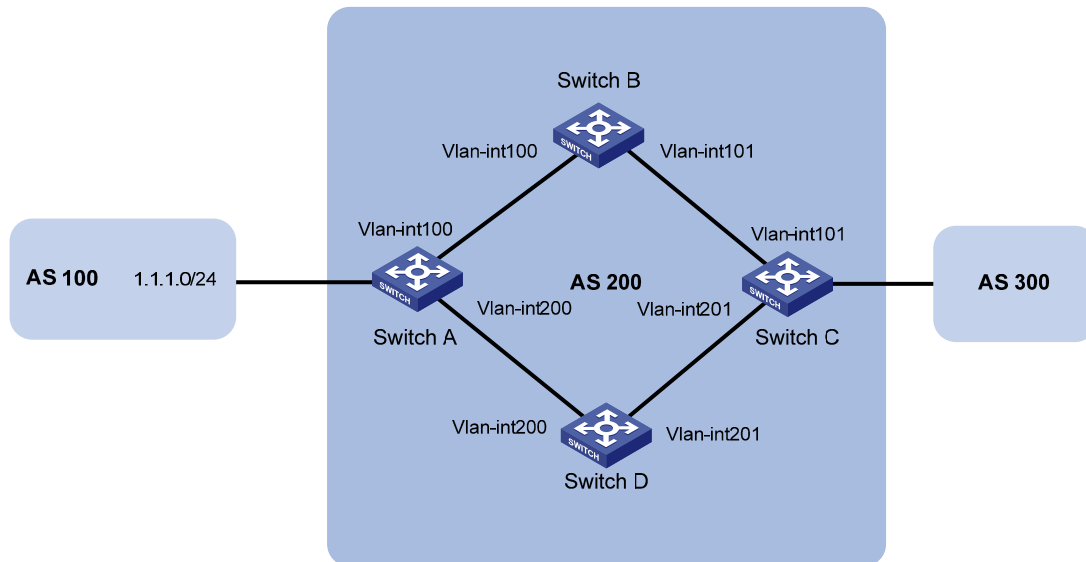
Configuring BFD for BGP

Network requirements

As shown in [Figure 98](#),

- Configure OSPF as the IGP in AS 200.
- Establish two IBGP connections between Switch A and Switch C. When both links are working, Switch C adopts the link Switch A ↔ Switch B ↔ Switch C to exchange packets with network 1.1.1.0/24. Configure BFD over the link. Then if the link fails, BFD can quickly detect the failure and notify it to BGP. Then the link Switch A ↔ Switch D ↔ Switch C takes effect immediately.

Figure 98 Network diagram



Device	Interface	IP address	Device	Interface	IP address
Switch A	Vlan-int100	3.0.1.1/24	Switch C	Vlan-int101	3.0.2.2/24
	Vlan-int200	2.0.1.1/24		Vlan-int201	2.0.2.2/24
Switch B	Vlan-int100	3.0.1.2/24	Switch D	Vlan-int200	2.0.1.2/24
	Vlan-int101	3.0.2.1/24		Vlan-int201	2.0.2.1/24

Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure OSPF to make sure that Switch A and Switch C are reachable to each other. (Details not shown.)
3. Configure BGP on Switch A:

Establish two IBGP connections between Switch A and Switch C.

```
<SwitchA> system-view
[SwitchA] bgp 100
[SwitchA-bgp] peer 3.0.2.2 as-number 200
[SwitchA-bgp] peer 2.0.2.2 as-number 200
[SwitchA-bgp] quit
```

When the two links between Switch A and Switch C are both up, Switch C adopts the link Switch A<—>Switch B<—>Switch C to exchange packets with network 1.1.1.0/24. (Set a higher MED value for route 1.1.1.0/24 sent to peer 2.0.2.2 on Switch A.)

- Create ACL 2000 to permit 1.1.1.0/24 to pass.


```
[SwitchA] acl number 2000
[SwitchA-acl-basic-2000] rule permit source 1.1.1.0 24
[SwitchA-acl-basic-2000] quit
```
- Create two route policies, **apply_med_50** and **apply_med_100**. Policy **apply_med_50** sets the MED for route 1.1.1.0/24 to 50. Policy **apply_med_100** sets that to 100.


```
[SwitchA] route-policy apply_med_50 permit node 10
[SwitchA-route-policy] if-match acl 2000
[SwitchA-route-policy] apply cost 50
[SwitchA-route-policy] quit
```

```
[SwitchA] route-policy apply_med_100 permit node 10
[SwitchA-route-policy] if-match acl 2000
[SwitchA-route-policy] apply cost 100
[SwitchA-route-policy] quit
```

- Apply routing policy **apply_med_50** to routes outgoing to peer 3.0.2.2, and apply routing policy **apply_med_100** to routes outgoing to peer 2.0.2.2.

```
[SwitchA] bgp 100
[SwitchA-bgp] peer 3.0.2.2 route-policy apply_med_50 export
[SwitchA-bgp] peer 2.0.2.2 route-policy apply_med_100 export
```

Configure BFD over the link to peer 3.0.2.2 so that when the link Switch A<—>Switch B<—>Switch C fails, BFD can quickly detect the failure and notify it to BGP, and then the link Switch A<—>Switch D<—>Switch C takes effect immediately.

```
[SwitchA-bgp] peer 3.0.2.2 bfd
[SwitchA-bgp] quit
```

4. Configure BGP on Switch C:

```
<SwitchC> system-view
[SwitchC] bgp 100
[SwitchC-bgp] peer 3.0.1.1 as-number 200
[SwitchC-bgp] peer 3.0.1.1 bfd
[SwitchC-bgp] peer 2.0.1.1 as-number 200
[SwitchC-bgp] quit
```

5. Configure BFD parameters (you can use default BFD parameters instead):

Configure Switch A.

```
[SwitchA] bfd session init-mode active
[SwitchA] interface vlan-interface 100
```

- Configure the minimum interval for transmitting BFD control packets as 500 milliseconds.


```
[SwitchA-Vlan-interface100] bfd min-transmit-interval 500
```
- Configure the minimum interval for receiving BFD control packets as 500 milliseconds.


```
[SwitchA-Vlan-interface100] bfd min-receive-interval 500
```
- Configure the detect multiplier as 7.


```
[SwitchA-Vlan-interface100] bfd detect-multiplier 7
```
- Configure the BFD authentication mode as plain-text authentication, and set the authentication key to **ibgpbfd**.

```
[SwitchA-Vlan-interface100] bfd authentication-mode simple 1 ibgpbfd
[SwitchA-Vlan-interface100] quit
```

Configure Switch C.

```
[SwitchC] bfd session init-mode active
[SwitchC] interface vlan-interface 101
[SwitchC-Vlan-interface101] bfd min-transmit-interval 500
[SwitchC-Vlan-interface101] bfd min-receive-interval 500
[SwitchC-Vlan-interface101] bfd detect-multiplier 7
[SwitchC-Vlan-interface101] bfd authentication-mode simple 1 ibgpbfd
[SwitchC-Vlan-interface101] return
```

6. Verify the configuration:

The following operations are made on Switch C. Operations on Switch A are similar. (Details not shown.)

Display detailed BFD session information.

<SwitchC> display bfd session verbose

Total Session Num: 1 Init Mode: Active

IP Session Working Under Ctrl Mode:

```
Local Discr: 17                              Remote Discr: 13
Source IP: 3.0.2.2                          Destination IP: 3.0.1.1
Session State: Up                            Interface: Vlan-interface101
Min Trans Inter: 500ms                      Act Trans Inter: 500ms
Min Recv Inter: 500ms                      Act Detect Inter: 3000ms
Recv Pkt Num: 57                            Send Pkt Num: 53
Hold Time: 2200ms                          Connect Type: Indirect
Running Up for: 00:00:06                    Auth mode: Simple
Protocol: BGP6
Diag Info: No Diagnostic
```

The output shows that a BFD session is established between Switch A's VLAN-interface 100 and Switch C's VLAN-interface 101 and that BFD runs properly.

Display BGP peer information on Switch C, and you can see that Switch C has established two BGP neighborships with Switch A.

<SwitchC> display bgp peer

BGP local router ID : 1.1.1.1

Local AS number : 200

Total number of peers : 2

Peers in established state : 2

Peer	AS	MsgRcvd	MsgSent	OutQ	PrefRcv	Up/Down	State
2.0.1.1	200	7	10	0	0	00:01:05	Established
3.0.1.1	200	7	10	0	0	00:01:34	Established

Display route 1.1.1.0/24 on Switch C, and you can see that Switch A and Switch C communicate through Switch B.

<SwitchC> display ip routing-table 1.1.1.0 24 verbose

Routing Table : Public

Summary Count : 2

```
Destination: 1.1.1.0/24
Protocol: BGP                              Process ID: 0
Preference: 0                              Cost: 50
NextHop: 3.0.1.1                          Interface: Vlan-interface101
BkNextHop: 0.0.0.0                        BkInterface:
RelyNextHop: 3.0.2.1                      Neighbor : 3.0.1.1
Tunnel ID: 0x0                            Label: NULL
State: Active Adv                        Age: 00h08m54s
Tag: 0
```

Destination: 1.1.1.0/24

```

    Protocol: BGP                Process ID: 0
Preference: 0                    Cost: 100
    NextHop: 2.0.1.1            Interface: Vlan-interface201
    BkNextHop: 0.0.0.0          BkInterface:
    RelyNextHop: 2.0.2.1        Neighbor : 2.0.1.1
    Tunnel ID: 0x0              Label: NULL
    State: Invalid Adv           Age: 00h08m54s
    Tag: 0

```

The output shows that Switch C has two routes to reach network 1.1.1.0/24: Switch C \leftrightarrow Switch B \leftrightarrow Switch A, which is the active route; Switch C \leftrightarrow Switch D \leftrightarrow Switch A, which is the backup route.

Enable BFD debugging on Switch C.

```

<SwitchC> debugging bfd scm
<SwitchC> debugging bfd event
<SwitchC> debugging bgp bfd
<SwitchC> terminal monitor
<SwitchC> terminal debugging

```

The following debugging information shows that: when the link between Switch A and Switch B fails, Switch C can quickly detect the link failure.

```

%Nov 5 11:42:24:172 2009 SwitchC BFD/5/BFD_CHANGE_FSM: Sess[3.0.2.2/3.0.1.1,
13/17,VLAN101,Ctrl], Sta: UP->DOWN, Diag: 1
%Nov 5 11:42:24:172 2009 SwitchC BGP/5/BGP_STATE_CHANGED: 3.0.1.1 state is changed
from ESTABLISHED to IDLE.
*Nov 5 11:42:24:187 2009 SwitchC RM/6/RMDEBUG: BGP_BFD: Recv BFD DOWN msg, Src IP
3.0.2.2, Dst IP 3.0.1.1, Instance ID 0.
*Nov 5 11:42:24:187 2009 SwitchC RM/6/RMDEBUG: BGP_BFD: Reset BGP session 3.0.1.1
for BFD session down.
*Nov 5 11:42:24:187 2009 SwitchC RM/6/RMDEBUG: BGP_BFD: Send DELETE msg to BFD,
Connection type DIRECT, Src IP 3.0.2.2, Dst IP 3.0.1.1, Instance ID 0.

```

Display route 1.1.1.0/24 on Switch C, and you can see that Switch A and Switch C communicate through Switch D.

```

<SwitchC> display ip routing-table 1.1.1.0 24 verbose
Routing Table : Public
Summary Count : 1

```

```

Destination: 1.1.1.0/24
    Protocol: BGP                Process ID: 0
Preference: 0                    Cost: 100
    NextHop: 2.0.1.1            Interface: Vlan-interface201
    BkNextHop: 0.0.0.0          BkInterface:
    RelyNextHop: 2.0.2.1        Neighbor : 2.0.1.1
    Tunnel ID: 0x0              Label: NULL
    State: Active Adv           Age: 00h09m54s
    Tag: 0

```

The output shows that Switch C has one route Switch C \leftrightarrow Switch D \leftrightarrow Switch A to reach network 1.1.1.0/24.

Troubleshooting BGP

BGP peer relationship not established

Symptom

Display BGP peer information by using the **display bgp peer** command. The state of the connection to a peer cannot become established.

Analysis

To become BGP peers, any two routers must establish a TCP session using port 179 and exchange Open messages successfully.

Solution

1. Use the **display current-configuration** command to check that the peer's AS number is correct.
2. Use the **display bgp peer command** to check that the peer's IP address is correct.
3. If a loopback interface is used, check that the loopback interface is specified with the **peer connect-interface** command.
4. If the peer is a non-direct EBGP peer, check that the **peer ebgp-max-hop** command is configured.
5. Check that a valid route to the peer is available.
6. Use the **ping** command to check the connectivity to the peer.
7. Use the **display tcp status** command to check the TCP connection.
8. Check whether an ACL disabling TCP port 179 is configured.

Configuring IPv6 static routing

Hardware compatibility

The HP 5500 SI Switch Series does not support VPN-related parameters.

Overview

Static routes are manually configured. They work well in simple networks. Proper configuration and use can improve network performance and ensure enough bandwidth for important applications.

However, static routes also have limitations. Any topology changes require manual configuration and modification to the relevant static routes.

The term "router" in this chapter refers to both routers and Layer 3 switches.

IPv6 static routes features

Similar to IPv4 static routes, IPv6 static routes work well in simple IPv6 network environments.

Their major difference lies in the destination and next hop addresses. IPv6 static routes use IPv6 addresses, whereas IPv4 static routes use IPv4 addresses.

Default IPv6 route

An IPv6 static route with a destination prefix of `::/0` is a default IPv6 route. The default route is used to forward packets that match no specific routes in the routing table.

Configuring an IPv6 static route

In small IPv6 networks, IPv6 static routes can be used to forward packets. In comparison to dynamic routes, it helps to save network bandwidth.

Before you configure an IPv6 static route, complete the following tasks:

- Configure parameters for the related interfaces.
- Configure link layer attributes for the related interfaces.
- Enable IPv6 packet forwarding.
- Make sure that the neighboring nodes can reach each other.

To configure an IPv6 static route:

Step	Command	Remarks
1.	Enter system view. <code>system-view</code>	N/A

Step	Command	Remarks
2. Configure an IPv6 static route.	<ul style="list-style-type: none"> Approach 1: ipv6 route-static <i>ipv6-address prefix-length</i> { <i>interface-type interface-number</i> [<i>next-hop-address</i>] <i>next-hop-address</i> vpn-instance <i>d-vpn-instance-name</i> <i>nexthop-address</i> } [preference <i>preference-value</i>] Approach 2: ipv6 route-static vpn-instance <i>s-vpn-instance-name</i> <1-6> <i>ipv6-address</i> <i>prefix-length</i> { <i>interface-type</i> <i>interface-number</i> [<i>next-hop-address</i>] <i>nexthop-address</i> [public] vpn-instance <i>d-vpn-instance-name nexthop-address</i> } [preference <i>preference-value</i>] 	Use either approach. The default preference of IPv6 static routes is 60.

NOTE:

If you specify a broadcast interface, such as an Ethernet interface or a VLAN interface, as the output interface for a static route, you must specify the next hop address.

Displaying and maintaining IPv6 static routes

Task	Command	Remarks
Display IPv6 static route information.	display ipv6 routing-table protocol static [inactive verbose] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Remove all IPv6 static routes.	delete ipv6 [vpn-instance <i>vpn-instance-name</i>] static-routes all	Available in system view

To delete a single IPv6 static route, use the **undo ipv6 route-static** command. To delete all IPv6 static routes, including the default route, use the **delete ipv6 static-routes all** command.

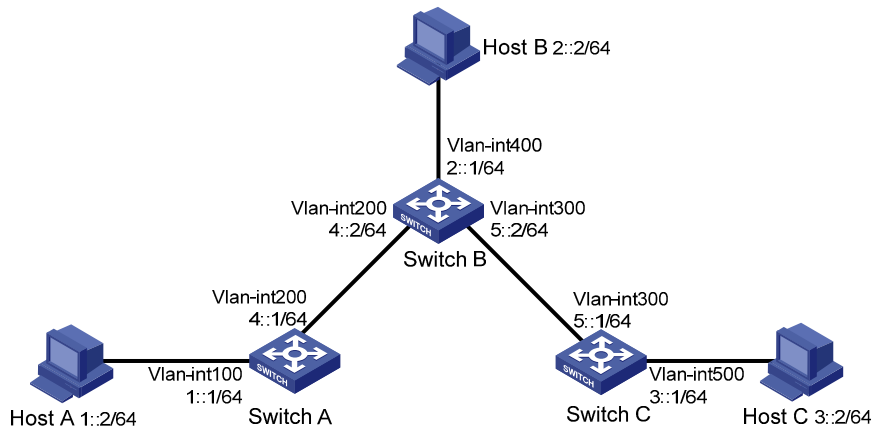
For more information about the **display ipv6 routing-table protocol static** [**inactive** | **verbose**] [[{ **begin** | **exclude** | **include** } *regular-expression*] command, see *Layer 3—IP Routing Command Reference*.

IPv6 static routing configuration example

Network requirements

As shown in [Figure 99](#), configure IPv6 static routes so that hosts can reach one another.

Figure 99 Network diagram



Configuration procedure

1. Configure the IPv6 addresses for all VLAN interfaces. (Details not shown.)
2. Configure IPv6 static routes:

Configure a default IPv6 static route on Switch A.

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] ipv6 route-static :: 0 4::2
```

Configure two IPv6 static routes on Switch B.

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] ipv6 route-static 1:: 64 4::1
[SwitchB] ipv6 route-static 3:: 64 5::1
```

Configure a default IPv6 static route on Switch C.

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] ipv6 route-static :: 0 5::2
```

3. Configure the IPv6 addresses and gateways for hosts:

Configure the IPv6 addresses for all the hosts based on the network diagram, configure the default gateway of Host A as 1::1, Host B as 2::1, and Host C as 3::1.

4. Verify the configuration:

Display the IPv6 routing table of Switch A.

```
[SwitchA] display ipv6 routing-table
```

Routing Table :

Destinations : 5 Routes : 5

Destination	: ::	Protocol	: Static
NextHop	: 4::2	Preference	: 60
Interface	: Vlan-interface200	Cost	: 0

Destination	: ::1/128	Protocol	: Direct
NextHop	: ::1	Preference	: 0
Interface	: InLoop0	Cost	: 0

Destination	: 1::/64	Protocol	: Direct
NextHop	: 1::1	Preference	: 0
Interface	: Vlan-interface100	Cost	: 0
Destination	: 1::1/128	Protocol	: Direct
NextHop	: ::1	Preference	: 0
Interface	: InLoop0	Cost	: 0
Destination	: FE80::/10	Protocol	: Direct
NextHop	: ::	Preference	: 0
Interface	: NULL0	Cost	: 0

Verify the connectivity with the **ping** command.

```
[SwitchA] ping ipv6 3::1
PING 3::1 : 56 data bytes, press CTRL_C to break
  Reply from 3::1
    bytes=56 Sequence=1 hop limit=254  time = 63 ms
  Reply from 3::1
    bytes=56 Sequence=2 hop limit=254  time = 62 ms
  Reply from 3::1
    bytes=56 Sequence=3 hop limit=254  time = 62 ms
  Reply from 3::1
    bytes=56 Sequence=4 hop limit=254  time = 63 ms
  Reply from 3::1
    bytes=56 Sequence=5 hop limit=254  time = 63 ms

--- 3::1 ping statistics ---
  5 packet(s) transmitted
  5 packet(s) received
  0.00% packet loss
  round-trip min/avg/max = 62/62/63 ms
```

Configuring RIPng

Hardware compatibility

The HP 5500 SI Switch Series does not support VPN-related parameters.

Introduction to RIPng

RIP next generation (RIPng) is an extension of RIP-2 for IPv4. Most RIP concepts are applicable in RIPng. The term "router" in this chapter refers to both routers and Layer 3 switches.

RIPng for IPv6 has the following basic differences from RIP:

- **UDP port number**—RIPng uses UDP port 521 for sending and receiving routing information.
- **Multicast address**—RIPng uses FF02:9 as the link-local-router multicast address.
- **Destination Prefix**—128-bit destination address prefix.
- **Next hop**—128-bit IPv6 address.
- **Source address**—RIPng uses FE80::/10 as the link-local source address.

RIPng working mechanism

RIPng is a routing protocol based on the distance vector (D-V) algorithm. RIPng uses UDP packets to exchange routing information through port 521.

RIPng uses a hop count to measure the distance to a destination. The hop count is the metric or cost. The hop count from a router to a directly connected network is 0. The hop count between two directly connected routers is 1. When the hop count is greater than or equal to 16, the destination network or host is unreachable.

By default, the routing update is sent every 30 seconds. If the router receives no routing updates from a neighbor within 180 seconds, the routes learned from the neighbor are considered unreachable. If no routing update is received within another 240 seconds, the router removes these routes from the routing table.

RIPng supports split horizon and poison reverse to prevent routing loops and route redistribution.

Each RIPng router maintains a routing database, which includes route entries of all reachable destinations. A route entry contains the following information:

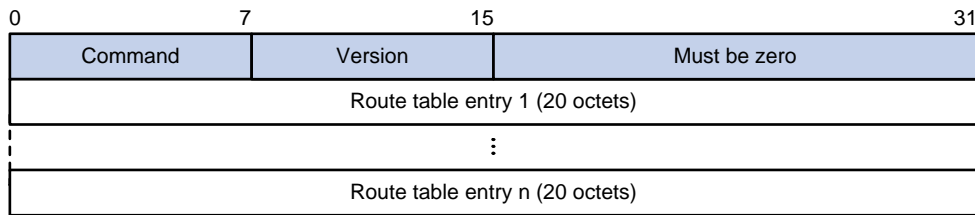
- **Destination address**—IPv6 address of a host or a network.
- **Next hop address**—IPv6 address of a neighbor along the path to the destination.
- **Egress interface**—Outbound interface that forwards IPv6 packets.
- **Metric**—Cost from the local router to the destination.
- **Route time**—Time elapsed since a route entry is last changed. Each time a route entry is modified, the routing time is set to 0.
- **Route tag**—Identifies the route used in a routing policy to control routing information. For more information about routing policy, see "[Configuring routing policies](#)."

RIPng packet format

Basic format

A RIPng packet consists of a header and multiple route table entries (RTEs). The maximum number of RTEs in a packet depends on the IPv6 MTU of the sending interface.

Figure 100 RIPng basic packet format



Packet header description:

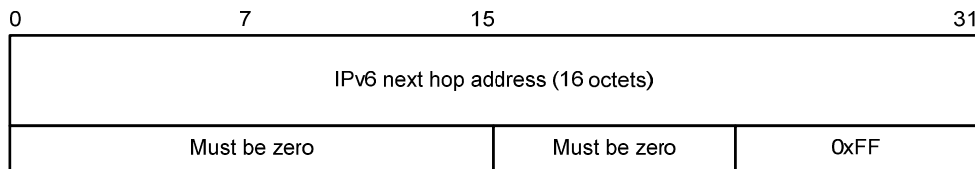
- **Command**—Type of message. 0x01 indicates Request; 0x02 indicates Response.
- **Version**—Version of RIPng. It can only be 0x01.
- **RTE**—Route table entry. It is 20 bytes for each entry.

RTE format

The following are types of RTEs in RIPng:

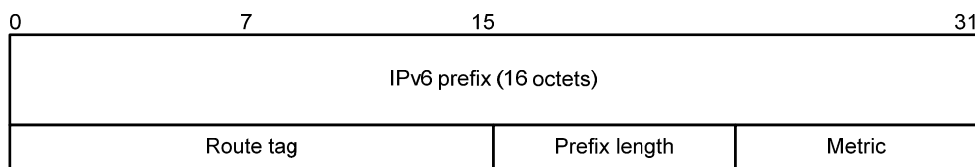
- **Next hop RTE**—Defines the IPv6 address of a next hop
- **IPv6 prefix RTE**—Describes the destination IPv6 address, route tag, prefix length and metric in the RIPng routing table.

Figure 101 Next hop RTE format



IPv6 next hop address is the IPv6 address of the next hop.

Figure 102 IPv6 prefix RTE format



IPv6 prefix RTE field description:

- **IPv6 prefix**—Destination IPv6 address prefix.
- **Route tag**—Route tag.
- **Prefix len**—Length of the IPv6 address prefix.
- **Metric**—Cost of a route.

RIPng packet processing procedure

Request packet

When a RIPng router first starts or needs to update entries in its routing table, usually a multicast request packet is sent to ask for needed routes from neighbors.

The receiving RIPng router processes RTEs in the request. If only one RTE exists with the IPv6 prefix and prefix length both being 0, and with a metric value of 16, the RIPng router will respond with the entire routing table information in response messages. If multiple RTEs exist in the request message, the RIPng router will examine each RTE, update its metric, and send the requested routing information to the requesting router in the response packet.

Response packet

The response packet containing the local routing table information is generated as follows:

- A response to a request
- An update periodically
- A triggered update caused by route change

After a router receives a response, it checks the validity of the response before adding the route to its routing table, such as whether the source IPv6 address is the link-local address and whether the port number is correct. The response packet that failed the check is discarded.

Protocols and standards

- RFC 2080, *RIPng for IPv6*
- RFC 2081, *RIPng Protocol Applicability Statement*

RIPng configuration task list

Task	Remarks
Configuring RIPng basic functions	Required
Configuring an additional routing metric	Optional
Configuring RIPng route summarization	Optional
Advertising a default route	Optional
Configuring RIPng route control	Configuring a RIPng route filtering policy
	Configuring a priority for RIPng
	Configuring RIPng route redistribution
	Configuring RIPng timers
Tuning and optimizing the RIPng network	Configuring split horizon and poison reverse
	Configuring zero field check on RIPng packets
	Configuring the maximum number of ECMP routes
Applying IPsec policies for RIPng	Optional

Configuring RIPng basic functions

This section presents the information to configure the basic RIPng features.

You must enable RIPng first before configuring other tasks, but it is not necessary for RIPng-related interface configurations, such as assigning an IPv6 address.

Before you configure RIPng basic functions, complete the following tasks:

- Enable IPv6 packet forwarding.
- Configure an IP address for each interface, and make sure that all neighboring nodes can reach each other.

To configure the basic RIPng functions:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a RIPng process and enter RIPng view.	ripng [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	Not created by default.
3. Return to system view.	quit	N/A
4. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
5. Enable RIPng on the interface.	ripng <i>process-id</i> enable	Disabled by default. If RIPng is not enabled on an interface, the interface will not send or receive a RIPng route.

Configuring RIPng route control

Before you configure RIPng, complete the following tasks:

- Configure an IPv6 address on each interface, and make sure that all neighboring nodes can reach each other.
- Configure RIPng basic functions.
- Define an IPv6 ACL before using it for route filtering. For related information, see *ACL and QoS Configuration Guide*.
- Define an IPv6 address prefix list before using it for route filtering. For related information, see "[Configuring routing policies](#)."

Configuring an additional routing metric

An additional routing metric can be added to the metric of an inbound or outbound RIP route.

The outbound additional metric is added to the metric of a sent route. The route's metric in the routing table is not changed.

The inbound additional metric is added to the metric of a received route before the route is added into the routing table, so the route's metric is changed.

To configure an inbound or outbound additional routing metric:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type interface-number</i>	N/A
3. Specify an inbound routing additional metric.	ripng metricin <i>value</i>	Optional. 0 by default.
4. Specify an outbound routing additional metric.	ripng metricout <i>value</i>	Optional. 1 by default.

Configuring RIPng route summarization

Step	Command
1. Enter system view.	system-view
2. Enter interface view.	interface <i>interface-type interface-number</i>
3. Advertise a summary IPv6 prefix.	ripng summary-address <i>ipv6-address prefix-length</i>

Advertising a default route

With this feature enabled, a default route is advertised through the specified interface regardless of whether the default route is available in the local IPv6 routing table.

To advertise a default route:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type interface-number</i>	N/A
3. Advertise a default route.	ripng default-route { only originate } [cost <i>cost</i>]	Not advertised by default.

Configuring a RIPng route filtering policy

Reference a configured IPv6 ACL or prefix list to filter received or advertised routing information. You can also filter outbound routes redistributed from a routing specific routing protocol.

To configure a RIPng route filtering policy:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIPng view.	ripng [<i>process-id</i>]	N/A
3. Configure a filter policy to filter incoming routes.	filter-policy { <i>acl6-number</i> ipv6-prefix <i>ipv6-prefix-name</i> } import	By default, RIPng does not filter incoming routing information.

Step	Command	Remarks
4. Configure a filter policy to filter outgoing routes.	filter-policy { <i>acl6-number</i> ipv6-prefix <i>ipv6-prefix-name</i> } export [<i>protocol</i> [<i>process-id</i>]]	By default, RIPng does not filter outgoing routing information.

Configuring a priority for RIPng

Any routing protocol has its own protocol priority used for optimal route selection. You can set a priority for RIPng manually. The smaller the value, the higher the priority.

To configure a RIPng priority:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIPng view.	ripng [<i>process-id</i>]	N/A
3. Configure a RIPng priority.	preference [route-policy <i>route-policy-name</i>] <i>preference</i>	Optional. By default, the RIPng priority is 100.

Configuring RIPng route redistribution

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIPng view.	ripng [<i>process-id</i>]	N/A
3. Configure a default routing metric for redistributed routes.	default cost <i>cost</i>	Optional. The default metric of redistributed routes is 0.
4. Redistribute routes from another routing protocol.	import-route <i>protocol</i> [<i>process-id</i>] [allow-ibgp] [cost <i>cost</i> route-policy <i>route-policy-name</i>] *	By default, no route redistribution is configured.

Tuning and optimizing the RIPng network

This section describes how to tune and optimize the performance of the RIPng network, as well as applications under special network environments. Before tuning and optimizing the RIPng network, complete the following tasks:

- Configure a network layer address for each interface.
- Configure the basic RIPng functions.

Configuring RIPng timers

Adjust RIPng timers to optimize the performance of the RIPng network. When adjusting RIPng timers, consider the network performance, and perform unified configurations on routers running RIPng to avoid unnecessary network traffic increase or route oscillation.

To configure RIPng timers:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIPng view.	ripng [<i>process-id</i>]	N/A
3. Configure RIPng timers.	timers { garbage-collect <i>garbage-collect-value</i> suppress <i>suppress-value</i> timeout <i>timeout-value</i> update <i>update-value</i> } *	Optional. The RIPng timers have the following defaults: <ul style="list-style-type: none"> • 30 seconds for the update timer • 180 seconds for the timeout timer • 120 seconds for the suppress timer • 120 seconds for the garbage-collect timer

Configuring split horizon and poison reverse

If both split horizon and poison reverse are configured, only the poison reverse function takes effect.

Configuring split horizon

The split horizon function disables a route learned from an interface from being advertised through the same interface to prevent routing loops between neighbors. HP recommends enabling split horizon to prevent routing loops.

To configure split horizon:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable the split horizon function.	ripng split-horizon	Optional. Enabled by default.

Configuring the poison reverse function

The poison reverse function enables a route learned from an interface to be advertised through the interface. However, the metric of the route is set to 16 (unreachable).

To configure poison reverse:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable the poison reverse function.	ripng poison-reverse	Disabled by default.

Configuring zero field check on RIPng packets

Some fields in the RIPng packet must be zero, which are called "zero fields". With zero field check on RIPng packets enabled, if such a field contains a non-zero value, the entire RIPng packet is discarded. If you are sure that all packets are trustworthy, disable the zero field check to reduce the CPU processing time.

To configure RIPng zero field check:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIPng view.	ripng [<i>process-id</i>]	N/A
3. Enable the zero field check.	checkzero	Optional. Enabled by default.

Configuring the maximum number of ECMP routes

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIPng view.	ripng [<i>process-id</i>]	N/A
3. Configure the maximum number of ECMP routes for load balancing.	maximum load-balancing <i>number</i>	Optional. 8 by default.

Applying IPsec policies for RIPng

To protect routing information and defend attacks, RIPng supports using an IPsec policy to authenticate protocol packets.

Outbound RIPng packets carry the Security Parameter Index (SPI) defined in the relevant IPsec policy. A device uses the SPI carried in a received packet to match against the configured IPsec policy. If they match, the device accepts the packet; otherwise, it discards the packet and will not establish a neighbor relationship with the sending device.

You can configure an IPsec policy for a RIPng process or interface. The IPsec policy configured for a process applies to all packets in the process. The IPsec policy configured on an interface applies to packets on the interface. If an interface and its process each have an IPsec policy configured, the interface uses its own IPsec policy.

Configuration prerequisites

Before you apply an IPsec policy for RIPng, complete following tasks:

- Create an IPsec proposal.
- Create an IPsec policy.

For more information about IPsec policy configuration, see *Security Configuration Guide*.

Configuration guidelines

An IPsec policy used for RIPng can only be in manual mode. For more information, see *Security Configuration Guide*.

Configuration procedure

To apply an IPsec policy in a process:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RIPng view.	ripng [<i>process-id</i>]	N/A
3. Apply an IPsec policy in the process.	enable ipsec-policy <i>policy-name</i>	Not configured by default.

To apply an IPsec policy on an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Apply an IPsec policy on the interface.	ripng ipsec-policy <i>policy-name</i>	Not configured by default.

Displaying and maintaining RIPng

Task	Command	Remarks
Display configuration information of a RIPng process.	display ripng [<i>process-id</i> vpn-instance <i>vpn-instance-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display routes in the RIPng database.	display ripng process-id database [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the routing information of a specified RIPng process.	display ripng process-id route [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display RIPng interface information.	display ripng process-id interface [<i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Reset a RIPng process.	reset ripng process-id process	Available in user view
Clear statistics of a RIPng process.	reset ripng process-id statistics	Available in user view

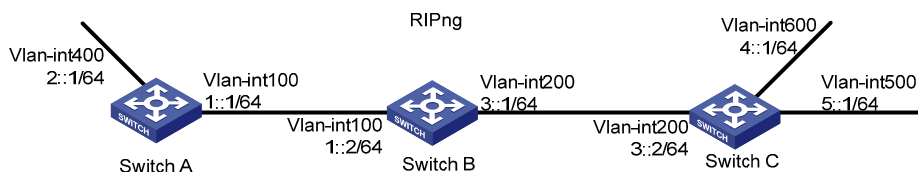
RIPng configuration examples

Configuring RIPng basic functions

Network requirements

In Figure 103, all switches run RIPng. Configure Switch B to filter the route (3::/64) learned from Switch C, which means the route will not be added to the routing table of Switch B, and Switch B will not forward it to Switch A.

Figure 103 Network diagram



Configuration procedure

1. Configure the IPv6 address for each interface. (Details not shown.)
2. Configure basic RIPng functions:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] ripng 1
[SwitchA-ripng-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ripng 1 enable
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 400
[SwitchA-Vlan-interface400] ripng 1 enable
[SwitchA-Vlan-interface400] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] ripng 1
[SwitchB-ripng-1] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] ripng 1 enable
[SwitchB-Vlan-interface200] quit
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ripng 1 enable
[SwitchB-Vlan-interface100] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] ripng 1
[SwitchC-ripng-1] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] ripng 1 enable
[SwitchC-Vlan-interface200] quit
```

```

[SwitchC] interface vlan-interface 500
[SwitchC-Vlan-interface500] ripng 1 enable
[SwitchC-Vlan-interface500] quit
[SwitchC] interface vlan-interface 600
[SwitchC-Vlan-interface600] ripng 1 enable
[SwitchC-Vlan-interface600] quit
# Display the routing table of Switch B.
[SwitchB] display ripng 1 route
    Route Flags: A - Aging, S - Suppressed, G - Garbage-collect
-----

Peer FE80::20F:E2FF:FE23:82F5 on Vlan-interface100
Dest 1::/64,
    via FE80::20F:E2FF:FE23:82F5, cost 1, tag 0, A, 6 Sec
Dest 2::/64,
    via FE80::20F:E2FF:FE23:82F5, cost 1, tag 0, A, 6 Sec

Peer FE80::20F:E2FF:FE00:100 on Vlan-interface200
Dest 3::/64,
    via FE80::20F:E2FF:FE00:100, cost 1, tag 0, A, 11 Sec
Dest 4::/64,
    via FE80::20F:E2FF:FE00:100, cost 1, tag 0, A, 11 Sec
Dest 5::/64,
    via FE80::20F:E2FF:FE00:100, cost 1, tag 0, A, 11 Sec
# Display the routing table of Switch A.
[SwitchA] display ripng 1 route
    Route Flags: A - Aging, S - Suppressed, G - Garbage-collect
-----

Peer FE80::200:2FF:FE64:8904 on Vlan-interface100
Dest 1::/64,
    via FE80::200:2FF:FE64:8904, cost 1, tag 0, A, 31 Sec
Dest 4::/64,
    via FE80::200:2FF:FE64:8904, cost 2, tag 0, A, 31 Sec
Dest 5::/64,
    via FE80::200:2FF:FE64:8904, cost 2, tag 0, A, 31 Sec
Dest 3::/64,
    via FE80::200:2FF:FE64:8904, cost 1, tag 0, A, 31 Sec

```

3. Configure Switch B to filter incoming and outgoing route:

```

[SwitchB] acl ipv6 number 2000
[SwitchB-acl6-basic-2000] rule deny source 3::/64
[SwitchB-acl6-basic-2000] rule permit
[SwitchB-acl6-basic-2000] quit
[SwitchB] ripng 1
[SwitchB-ripng-1] filter-policy 2000 import
[SwitchB-ripng-1] filter-policy 2000 export
# Display routing tables of Switch B and Switch A.
[SwitchB] display ripng 1 route

```



```

Route Flags: A - Aging, S - Suppressed, G - Garbage-collect
-----

Peer FE80::20F:E2FF:FE23:82F5 on Vlan-interface100
Dest 1::/64,
    via FE80::20F:E2FF:FE23:82F5, cost 1, tag 0, A, 2 Sec
Dest 2::/64,
    via FE80::20F:E2FF:FE23:82F5, cost 1, tag 0, A, 2 Sec

Peer FE80::20F:E2FF:FE00:100 on Vlan-interface200
Dest 4::/64,
    via FE80::20F:E2FF:FE00:100, cost 1, tag 0, A, 5 Sec
Dest 5::/64,
    via FE80::20F:E2FF:FE00:100, cost 1, tag 0, A, 5 Sec
[SwitchA] display ripng 1 route
Route Flags: A - Aging, S - Suppressed, G - Garbage-collect
-----

Peer FE80::20F:E2FF:FE00:1235 on Vlan-interface100
Dest 1::/64,
    via FE80::20F:E2FF:FE00:1235, cost 1, tag 0, A, 2 Sec
Dest 4::/64,
    via FE80::20F:E2FF:FE00:1235, cost 2, tag 0, A, 2 Sec
Dest 5::/64,
    via FE80::20F:E2FF:FE00:1235, cost 2, tag 0, A, 2 Sec

```

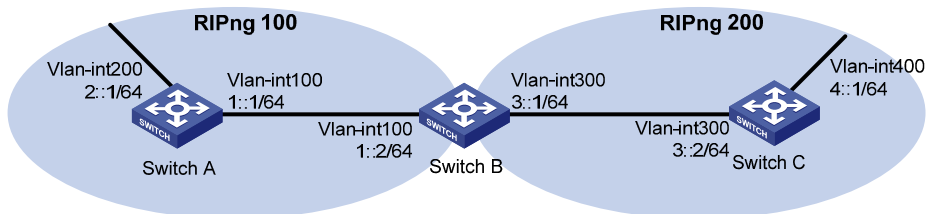
Configuring RIPng route redistribution

Network requirements

Two RIPng processes are running on Switch B, which communicates with Switch A through RIPng 100 and with Switch C through RIPng 200.

Configure route redistribution on Switch B, letting the two RIPng processes redistribute routes from each other. Set the default cost of redistributed routes from RIPng 200 to 3.

Figure 104 Network diagram



Configuration procedure

1. Configure IPv6 addresses for the interfaces. (Details not shown.)
2. Configure RIPng basic functions:
 - # Enable RIPng 100 on Switch A.
 - <SwitchA> system-view

```
[SwitchA] ripng 100
[SwitchA-ripng-100] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ripng 100 enable
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 200
[SwitchA-Vlan-interface200] ripng 100 enable
[SwitchA-Vlan-interface200] quit
```

Enable RIP 100 and RIP 200 on Switch B.

```
<SwitchB> system-view
[SwitchB] ripng 100
[SwitchB-ripng-100] quit
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ripng 100 enable
[SwitchB-Vlan-interface100] quit
[SwitchB] ripng 200
[SwitchB-ripng-200] quit
[SwitchB] interface vlan-interface 300
[SwitchB-Vlan-interface300] ripng 200 enable
[SwitchB-Vlan-interface300] quit
```

Enable RIPng 200 on Switch C.

```
<SwitchC> system-view
[SwitchC] ripng 200
[SwitchC] interface vlan-interface 300
[SwitchC-Vlan-interface300] ripng 200 enable
[SwitchC-Vlan-interface300] quit
[SwitchC] interface vlan-interface 400
[SwitchC-Vlan-interface400] ripng 200 enable
[SwitchC-Vlan-interface400] quit
```

Display the routing table of Switch A.

```
[SwitchA] display ipv6 routing-table
```

Routing Table :

Destinations : 6 Routes : 6

```
Destination: ::1/128                      Protocol : Direct
NextHop      : ::1                        Preference: 0
Interface    : InLoop0                    Cost      : 0
```

```
Destination: 1::/64                      Protocol : Direct
NextHop      : 1::1                        Preference: 0
Interface    : Vlan100                    Cost      : 0
```

```
Destination: 1::1/128                    Protocol : Direct
NextHop      : ::1                        Preference: 0
Interface    : InLoop0                    Cost      : 0
```

```
Destination: 2::/64                      Protocol : Direct
NextHop      : 2::1                        Preference: 0
```

```

Interface   : Vlan200                               Cost       : 0

Destination: 2::1/128                               Protocol   : Direct
NextHop     : ::1                                   Preference: 0
Interface   : InLoop0                               Cost       : 0

Destination: FE80::/10                              Protocol   : Direct
NextHop     : ::                                   Preference: 0
Interface   : NULL0                                 Cost       : 0

```

3. Configure RIPng route redistribution:

Configure route redistribution between the two RIPng processes on Switch B.

```

[SwitchB] ripng 100
[SwitchB-ripng-100] default cost 3
[SwitchB-ripng-100] import-route ripng 200
[SwitchB-ripng-100] quit
[SwitchB] ripng 200
[SwitchB-ripng-200] import-route ripng 100
[SwitchB-ripng-200] quit

```

Display the routing table of Switch A.

```

[SwitchA] display ipv6 routing-table
Routing Table :
                Destinations : 7           Routes : 7

```

```

Destination: ::1/128                               Protocol   : Direct
NextHop     : ::1                                   Preference: 0
Interface   : InLoop0                               Cost       : 0

Destination: 1::/64                                 Protocol   : Direct
NextHop     : 1::1                                  Preference: 0
Interface   : Vlan100                               Cost       : 0

Destination: 1::1/128                              Protocol   : Direct
NextHop     : ::1                                   Preference: 0
Interface   : InLoop0                               Cost       : 0

Destination: 2::/64                                 Protocol   : Direct
NextHop     : 2::1                                  Preference: 0
Interface   : Vlan200                               Cost       : 0

Destination: 2::1/128                              Protocol   : Direct
NextHop     : ::1                                   Preference: 0
Interface   : InLoop0                               Cost       : 0

Destination: 4::/64                                 Protocol   : RIPng
NextHop     : FE80::200:BFF:FE01:1C02              Preference: 100
Interface   : Vlan100                               Cost       : 4

Destination: FE80::/10                              Protocol   : Direct

```

NextHop : ::
Interface : NULL0

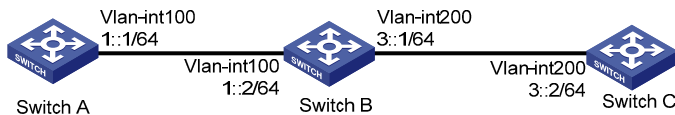
Preference: 0
Cost : 0d

Configuring RIPng IPsec policies

Network requirements

In the following figure, configure RIPng on the switches, and configure IPsec policies on the switches to authenticate and encrypt protocol packets.

Figure 105 Network diagram



Configuration procedure

1. Configure IPv6 addresses for interfaces. (Details not shown.)
2. Configure RIPng basic functions:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] ripng 1
[SwitchA-ripng-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ripng 1 enable
[SwitchA-Vlan-interface100] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] ripng 1
[SwitchB-ripng-1] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] ripng 1 enable
[SwitchB-Vlan-interface200] quit
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ripng 1 enable
[SwitchB-Vlan-interface100] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] ripng 1
[SwitchC-ripng-1] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] ripng 1 enable
[SwitchC-Vlan-interface200] quit
```

3. Configure RIPng IPsec policies:

On Switch A, create an IPsec proposal named **tran1**, and set the encapsulation mode to transport mode, the security protocol to ESP, the encryption algorithm to DES, and authentication algorithm to SHA1; create an IPsec policy named **policy001**, specify the manual mode for it,

reference IPsec proposal **tran1**, set the SPIs of the inbound and outbound SAs to 12345, and the keys for the inbound and outbound SAs using ESP to abcdefg.

```
[SwitchA] ipsec proposal tran1
[SwitchA-ipsec-proposal-tran1] encapsulation-mode transport
[SwitchA-ipsec-proposal-tran1] transform esp
[SwitchA-ipsec-proposal-tran1] esp encryption-algorithm des
[SwitchA-ipsec-proposal-tran1] esp authentication-algorithm sha1
[SwitchA-ipsec-proposal-tran1] quit
[SwitchA] ipsec policy policy001 10 manual
[SwitchA-ipsec-policy-manual-policy001-10] proposal tran1
[SwitchA-ipsec-policy-manual-policy001-10] sa spi outbound esp 12345
[SwitchA-ipsec-policy-manual-policy001-10] sa spi inbound esp 12345
[SwitchA-ipsec-policy-manual-policy001-10] sa string-key outbound esp abcdefg
[SwitchA-ipsec-policy-manual-policy001-10] sa string-key inbound esp abcdefg
[SwitchA-ipsec-policy-manual-policy001-10] quit
```

On Switch B, create an IPsec proposal named **tran1**, and set the encapsulation mode to transport mode, the security protocol to ESP, the encryption algorithm to DES, and authentication algorithm to SHA1; create an IPsec policy named **policy001**, specify the manual mode for it, reference IPsec proposal **tran1**, set the SPIs of the inbound and outbound SAs to 12345, and the keys for the inbound and outbound SAs using ESP to abcdefg.

```
[SwitchB] ipsec proposal tran1
[SwitchB-ipsec-proposal-tran1] encapsulation-mode transport
[SwitchB-ipsec-proposal-tran1] transform esp
[SwitchB-ipsec-proposal-tran1] esp encryption-algorithm des
[SwitchB-ipsec-proposal-tran1] esp authentication-algorithm sha1
[SwitchB-ipsec-proposal-tran1] quit
[SwitchB] ipsec policy policy001 10 manual
[SwitchB-ipsec-policy-manual-policy001-10] proposal tran1
[SwitchB-ipsec-policy-manual-policy001-10] sa spi outbound esp 12345
[SwitchB-ipsec-policy-manual-policy001-10] sa spi inbound esp 12345
[SwitchB-ipsec-policy-manual-policy001-10] sa string-key outbound esp abcdefg
[SwitchB-ipsec-policy-manual-policy001-10] sa string-key inbound esp abcdefg
[SwitchB-ipsec-policy-manual-policy001-10] quit
```

On Switch C, create an IPsec proposal named **tran1**, and set the encapsulation mode to transport mode, the security protocol to ESP, the encryption algorithm to DES, and authentication algorithm to SHA1; create an IPsec policy named **policy001**, specify the manual mode for it, reference IPsec proposal **tran1**, set the SPIs of the inbound and outbound SAs to 12345, and the keys for the inbound and outbound SAs using ESP to abcdefg.

```
[SwitchC] ipsec proposal tran1
[SwitchC-ipsec-proposal-tran1] encapsulation-mode transport
[SwitchC-ipsec-proposal-tran1] transform esp
[SwitchC-ipsec-proposal-tran1] esp encryption-algorithm des
[SwitchC-ipsec-proposal-tran1] esp authentication-algorithm sha1
[SwitchC-ipsec-proposal-tran1] quit
[SwitchC] ipsec policy policy001 10 manual
[SwitchC-ipsec-policy-manual-policy001-10] proposal tran1
[SwitchC-ipsec-policy-manual-policy001-10] sa spi outbound esp 12345
[SwitchC-ipsec-policy-manual-policy001-10] sa spi inbound esp 12345
```

```
[SwitchC-ipsec-policy-manual-policy001-10] sa string-key outbound esp abcdefg
[SwitchC-ipsec-policy-manual-policy001-10] sa string-key inbound esp abcdefg
[SwitchC-ipsec-policy-manual-policy001-10] quit
```

4. Apply the IPsec policies in the RIPng process:

Configure Switch A.

```
[SwitchA] ripng 1
[SwitchA-ripng-1] enable ipsec-policy policy001
[SwitchA-ripng-1] quit
```

Configure Switch B.

```
[SwitchB] ripng 1
[SwitchB-ripng-1] enable ipsec-policy policy001
[SwitchB-ripng-1] quit
```

Configure Switch C.

```
[SwitchC] ripng 1
[SwitchC-ripng-1] enable ipsec-policy policy001
[SwitchC-ripng-1] quit
```

5. Verify the configuration:

RIPng traffic between Switches A, B and C is protected by IPsec.

Configuring OSPFv3

Hardware compatibility

The HP 5500 SI Switch Series does not support OSPFv3.

Introduction to OSPFv3

OSPFv3 overview

Open Shortest Path First version 3 (OSPFv3) supports IPv6 and complies with RFC 2740 (OSPF for IPv6). The term "router" in this chapter refers to both routers and Layer 3 switches.

OSPFv3 and OSPFv2 have the following similarities:

- 32-bits router ID and area ID
- Packets, including Hello, DD (Data Description), LSR (Link State Request), LSU (Link State Update), and LSAck (Link State Acknowledgment)
- Mechanism for finding neighbors and establishing adjacencies
- Mechanism for LSA flooding and aging

OSPFv3 and OSPFv2 have the following differences:

- OSPFv3 runs on a per-link basis, and OSPFv2 runs on a per-IP-subnet basis.
- OSPFv3 supports multiple instances per link, but OSPFv2 does not.
- OSPFv3 identifies neighbors by Router ID, and OSPFv2 by IP address.

OSPFv3 packets

OSPFv3 has the following packet types: hello, DD, LSR, LSU, and LSAck. These packets have the same packet header, which is different from the OSPFv2 packet header. The OSPFv3 packet header is only 16 bytes in length, has no authentication field, and is added with an Instance ID field to support VPN per link.

Figure 106 OSPFv3 packet header

0	15	31
Version #	Type	Packet length
Router ID		
Area ID		
Checksum	Instance ID	0

Major fields for OSPFv3 packets are as follows:

- **Version #**—Version of OSPF, which is 3 for OSPFv3.
- **Type**—Type of OSPF packet; types 1 to 5 are hello, DD, LSR, LSU, and LSAck.
- **Packet length**—Packet length in bytes, including header.

- **Instance ID**—Instance ID for a link.
- **0**—Reserved. It must be 0.

OSPFv3 LSA types

OSPFv3 sends routing information in LSAs, which, as defined in RFC 2740, have the following types:

- **Router-LSA**—Originated by all routers. This LSA describes the collected states of the router's interfaces to an area, and is flooded throughout a single area only.
- **Network-LSA**—Originated for broadcast and NBMA networks by the Designated Router. This LSA contains the list of routers connected to the network, and is flooded throughout a single area only.
- **Inter-Area-Prefix-LSA**—Similar to Type 3 LSA of OSPFv2, originated by ABRs (Area Border Routers), and flooded throughout the LSA's associated area. Each Inter-Area-Prefix-LSA describes a route with IPv6 address prefix to a destination outside the area, yet still inside the AS (an inter-area route).
- **Inter-Area-Router-LSA**—Similar to Type 4 LSA of OSPFv2, originated by ABRs and flooded throughout the LSA's associated area. Each Inter-Area-Router-LSA describes a route to ASBR (Autonomous System Boundary Router).
- **AS-external-LSA**—Originated by ASBRs, and flooded throughout the AS (except Stub and NSSA areas). Each AS-external-LSA describes a route to another autonomous system. A default route can be described by an AS-external-LSA.
- **Link-LSA**—A router originates a separate Link-LSA for each attached link. Link-LSAs have link-local flooding scope. Each Link-LSA describes the IPv6 address prefix of the link and Link-local address of the router.
- **Intra-Area-Prefix-LSA**—Each Intra-Area-Prefix-LSA contains IPv6 prefix information on a router, stub area, or transit area information, and has area flooding scope. It was introduced because Router-LSAs and Network-LSAs do not contain address information.

RFC 5187 defines the Type 11 LSA, Grace-LSA. A Grace-LSA is generated by a GR (Graceful Restart) Restarter at reboot and transmitted on the local link. The restarter describes the cause and interval of the reboot in the Grace-LSA to tell its neighbors that it performs a GR operation.

OSPFv3 timers

Timers in OSPFv3 include the following:

- OSPFv3 packet timer
- LSA delay timer
- SPF timer
- GR timer

OSPFv3 packet timer

Hello packets are sent periodically between neighboring routers for finding and maintaining neighbor relationships, or for DR or BDR election. The hello interval must be identical on neighboring interfaces. The smaller the hello interval, the faster the network convergence speed and the bigger the network load.

If a router does not receive a hello packet from a neighbor within a given period—dead interval, it then declares the peer down.

After sending an LSA to its adjacency, a router waits for an acknowledgment from the adjacency. If no response is received after the retransmission interval elapses, the router will send the LSA again. The retransmission interval must be longer than the round-trip time of the LSA.

LSA delay time

Each LSA has an age in the local LSDB (incremented by one per second), but an LSA does not age on transmission. You must add an LSA delay time into the age time before transmission, which is important for low-speed networks.

SPF timer

Whenever the LSDB changes, an SPF calculation occurs. If recalculations become frequent, a large amount of resources will be occupied. You can adjust the SPF calculation interval and delay time to protect networks from being overloaded by frequent changes.

GR timer

If a failure to establish adjacencies occurs during a GR, the device will be in the GR process for a long time. To avoid this, configure the GR timer for the device to exit the GR process when the timer expires.

OSPFv3 features supported

- Basic features defined in RFC 2740
- OSPFv3 stub area
- OSPFv3 multi-process
- VPN instances
- OSPFv3 GR
- BFD

Protocols and standards

- RFC 2740, *OSPF for IPv6*
- RFC 2328, *OSPF Version 2*
- RFC 5187, *OSPFv3 Graceful Restart*

OSPFv3 configuration task list

Task	Remarks
Enabling OSPFv3	Required
Configuring OSPFv3 area parameters	Configuring an OSPFv3 stub area Optional
	Configuring an OSPFv3 virtual link Optional
Configuring OSPFv3 network types	Configuring the OSPFv3 network type for an interface Optional
	Configuring an NBMA or P2MP neighbor Optional
Configuring OSPFv3 routing information control	Configuring OSPFv3 route summarization Optional
	Configuring OSPFv3 inbound route filtering Optional
	Configuring an OSPFv3 cost for an interface Optional
	Configuring the maximum number of OSPFv3 ECMP routes Optional

Task	Remarks	
Tuning and optimizing OSPFv3 networks	Configuring a priority for OSPFv3	Optional
	Configuring OSPFv3 route redistribution	Optional
	Configuring OSPFv3 timers	Optional
	Configuring a DR priority for an interface	Optional
	Ignoring MTU check for DD packets	Optional
	Disabling interfaces from receiving and sending OSPFv3 packets	Optional
	Enabling the logging of neighbor state changes	Optional
Configuring OSPFv3 GR	Configuring GR Restarter	Optional
	Configuring GR Helper	Optional
Configuring BFD for OSPFv3	Optional	
Applying IPsec policies for OSPFv3	Optional	

Enabling OSPFv3

Configuration prerequisites

Before you enable OSPFv3, complete the following tasks:

- Make neighboring nodes accessible with each other at the network layer.
- Enable IPv6 packet forwarding.

Enabling OSPFv3

To enable an OSPFv3 process on a router, you must enable the OSPFv3 process globally, assign the OSPFv3 process a router ID, and enable the OSPFv3 process on related interfaces.

A router ID uniquely identifies a router within an AS. You must specify a unique router ID for each OSPFv3 router within the AS to ensure normal operation. If a router runs multiple OSPFv3 processes, you must specify a unique router ID for each process.

An OSPFv3 process ID has only local significance. Process 1 on a router can exchange packets with process 2 on another router.

To enable OSPFv3:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable an OSPFv3 process and enter its view.	ospfv3 [<i>process-id</i>] [vpn-instance <i>vpn-instance-name</i>]	By default, no OSPFv3 process is enabled.
3. Specify a router ID.	router-id <i>router-id</i>	N/A
4. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A

Step	Command	Remarks
5. Enable an OSPFv3 process on the interface.	ospfv3 <i>process-id</i> area <i>area-id</i> [instance <i>instance-id</i>]	Not enabled by default.

Configuring OSPFv3 area parameters

The stub area and virtual link features of OSPFv3 are the same as OSPFv2.

Splitting an OSPFv3 AS into multiple areas reduces the number of LSAs and extends OSPFv3 applications. For those non-backbone areas residing on the AS boundary, configure them as stub areas to further reduce the size of routing tables and the number of LSAs.

Non-backbone areas exchange routing information through the backbone area. The backbone and non-backbone areas—including the backbone itself—must be contiguous. In practice, necessary physical links may not be available for this connectivity. You can configure virtual links to address the problem.

Configuration prerequisites

Before you configure OSPFv3 area parameters, complete the following tasks:

- Enable IPv6 packet forwarding.
- Configure OSPFv3 basic functions.

Configuring an OSPFv3 stub area

Follow these guidelines when you configure an OSPFv3 stub area:

- You cannot remove an OSPFv3 area directly. The area can be removed only when you remove all configurations in area view and all interfaces attached to the area become down.
- All the routers attached to a stub area must be configured with the **stub** command. The keyword **no-summary** is only available on the ABR of the stub area.
- If you use the **stub** command with the keyword **no-summary** on an ABR, the ABR advertises a default route in an Inter-Area-Prefix-LSA into the stub area. No AS-external-LSA, Inter-Area-Prefix-LSA, or Inter-Area-Router-LSA is advertised in the area. The stub area of this kind is also known as a "totally stub area."

To configure an OSPFv3 stub area:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPFv3 view.	ospfv3 [<i>process-id</i>]	N/A
3. Enter OSPFv3 area view.	area <i>area-id</i>	N/A
4. Configure the area as a stub area.	stub [no-summary]	Not configured by default.
5. Specify a cost for the default route advertised to the stub area.	default-cost <i>value</i>	Optional. 1 by default.

Configuring an OSPFv3 virtual link

You can configure a virtual link to maintain connectivity between a non-backbone area and the backbone, or in the backbone itself.

ⓘ IMPORTANT:

- Both ends of a virtual link are ABRs that must be configured with the **vlink-peer** command.
- Do not configure virtual links in the areas of a GR-capable process.

To configure a virtual link:

Step	Command
1. Enter system view.	system-view
2. Enter OSPFv3 view.	ospfv3 [<i>process-id</i>]
3. Enter OSPFv3 area view.	area <i>area-id</i>
4. Configure a virtual link.	vlink-peer <i>router-id</i> [hello <i>seconds</i> retransmit <i>seconds</i> trans-delay <i>seconds</i> dead <i>seconds</i> instance <i>instance-id</i>] *

Configuring OSPFv3 network types

OSPFv3 classifies networks into the following types by the link layer protocol.

By default, the default OSPFv3 interface network types vary with the link layer protocols of the interfaces:

- When the link layer protocol is PPP, OSPFv3 considers the network type as P2P by default.
- When the link layer protocol is Ethernet, OSPFv3 considers the network type as broadcast by default.

You can change the network type of an OSPFv3 interface as needed. For example:

- An NBMA network must be fully connected. Any two routers in the network must be directly reachable to each other through a virtual circuit. If no such direct link is available, you must change the network type through a command.
- If direct connections are not available between some routers in an NBMA network, the type of interfaces associated must be configured as P2MP, or as P2P for interfaces with only one neighbor.

Configuration prerequisites

Before you configure OSPFv3 network types, complete the following tasks:

- Configure IPv6 functions.
- Configure OSPFv3 basic functions.

Configuring the OSPFv3 network type for an interface

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure a network type for the OSPFv3 interface.	ospfv3 network-type { broadcast nbma p2mp [non-broadcast] p2p } [instance <i>instance-id</i>]	Optional. The network type of an interface depends on the media type of the interface.

Configuring an NBMA or P2MP neighbor

For NBMA and P2MP interfaces (only when in unicast mode), specify the link-local IP addresses of their neighbors because these interfaces cannot find neighbors through broadcasting hello packets. You can also specify DR priorities for neighbors.

To configure an NBMA or P2MP (unicast) neighbor and its DR priority:

Step	Command
1. Enter system view.	system-view
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>
3. Specify an NBMA or P2MP (unicast) neighbor and its DR priority.	ospfv3 peer <i>ipv6-address</i> [dr-priority <i>dr-priority</i>] [instance <i>instance-id</i>]

Configuring OSPFv3 routing information control

This section describes how to configure the control of OSPF routing information advertisement and reception, and redistribution from other protocols.

Configuration prerequisites

Before you configure OSPFv3 routing information control, complete the following tasks:

- Enable IPv6 packet forwarding.
- Configure OSPFv3 basic functions.

Configuring OSPFv3 route summarization

If contiguous network segments exist in an area, use the **abr-summary** command to summarize them into one network segment on the ABR. The ABR will advertise only the summary route. Any LSA in the specified network segment will not be advertised, reducing the LSDB size in other areas.

To configure route summarization:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPFv3 view.	ospfv3 [<i>process-id</i>]	N/A
3. Enter OSPFv3 area view.	area <i>area-id</i>	N/A

Step	Command	Remarks
4. Configure a summary route.	abr-summary <i>ipv6-address</i> <i>prefix-length</i> [not-advertise]	Not configured by default. The abr-summary command takes effect on ABRs only.

Configuring OSPFv3 inbound route filtering

According to some rules, you can configure OSPFv3 to filter routes that are computed from received LSAs.

To configure OSPFv3 inbound route filtering:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPFv3 view.	ospfv3 [<i>process-id</i>]	N/A
3. Configure inbound route filtering.	filter-policy { <i>acl-number</i> ipv6-prefix <i>ipv6-prefix-name</i> } import	Not configured by default. The filter-policy import command can only filter routes computed by OSPFv3. Only routes not filtered out can be added into the local routing table.

Configuring an OSPFv3 cost for an interface

You can configure an OSPFv3 cost for an interface with one of the following methods:

- Configure the cost value in interface view.
- Configure a bandwidth reference value for the interface, and OSPFv3 computes the cost automatically based on the bandwidth reference value: Interface OSPFv3 cost = Bandwidth reference value (100 Mbps) ÷ Interface bandwidth (Mbps). If the calculated cost is greater than 65535, the value of 65535 is used; if the calculated cost is smaller than 1, the value of 1 is used.

If the cost value is not configured for an interface, OSPFv3 computes the interface cost value automatically.

To configure an OSPFv3 cost for an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure an OSPFv3 cost for the interface.	ospfv3 cost <i>value</i> [instance <i>instance-id</i>]	Optional. The default cost depends on the interface type: 1 for a VLAN interface; 0 for a loopback interface; computed according to the bandwidth for other interfaces.

To configure a bandwidth reference value:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPFv3 view.	ospfv3 [<i>process-id</i>]	N/A
3. Configure a bandwidth reference value.	bandwidth-reference <i>value</i>	Optional. 100 Mbps by default.

Configuring the maximum number of OSPFv3 ECMP routes

Perform this task to implement load sharing over ECMP routes.

To configure the maximum number of ECMP routes:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPFv3 view.	ospfv3 [<i>process-id</i>]	N/A
3. Specify the maximum number of ECMP routes.	maximum load-balancing <i>maximum</i>	Optional. 8 by default.

Configuring a priority for OSPFv3

A router may run multiple routing protocols. The system assigns a priority to each protocol. When these routing protocols find the same route, the route found by the protocol with the highest priority is selected.

To configure a priority for OSPFv3:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPFv3 view.	ospfv3 [<i>process-id</i>]	N/A
3. Configure a priority for OSPFv3.	preference [<i>ase</i>] [route-policy <i>route-policy-name</i>] <i>preference</i>	Optional. By default, the priority of OSPFv3 internal routes is 10, and priority of OSPFv3 external routes is 150.

Configuring OSPFv3 route redistribution

Follow these guidelines when you configure OSPFv3 route redistribution:

- Executing the **import-route** or **default-route-advertise** command on a router makes it become an ASBR.
- You can only inject and advertise a default route using the **default-route-advertise** command.
- Because OSPFv3 is a link state routing protocol, it cannot directly filter LSAs to be advertised. You must filter redistributed routes first. Routes that are not filtered out can be advertised in LSAs into the routing domain.

- Using the **filter-policy export** command filters routes redistributed with the **import-route** command. If the **import-route** command is not configured, executing the **filter-policy export** command does not take effect.

To configure OSPFv3 route redistribution:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPFv3 view.	ospfv3 [<i>process-id</i>]	N/A
3. Specify a default cost for redistributed routes.	default cost <i>value</i>	Optional. 1 by default.
4. Redistribute routes from another protocol, or another OSPFv3 process.	import-route <i>protocol</i> [<i>process-id</i> allow-ibgp] [cost <i>value</i> route-policy <i>route-policy-name</i> type <i>type</i>] *	Not configured by default.
5. Inject a default route.	default-route-advertise [always cost <i>value</i> type <i>type</i> route-policy <i>route-policy-name</i>] *	Optional. Not injected by default.
6. Filter redistributed routes.	filter-policy { <i>acl6-number</i> ipv6-prefix <i>ipv6-prefix-name</i> } export [isisv6 <i>process-id</i> ospfv3 <i>process-id</i> ripng <i>process-id</i> bgp4+ direct static]	Optional. Not configured by default.

Tuning and optimizing OSPFv3 networks

This section describes configurations of OSPFv3 timers, interface DR priority, MTU check ignorance for DD packets, and disabling interfaces from sending OSPFv3 packets.

The following are OSPFv3 timers:

- Packet timer**—Specified to adjust topology convergence speed and network load.
- LSA delay timer**—Specified especially for low-speed links.
- SPF timer**—Specified to protect networks from being overloaded due to frequent network changes.

For a broadcast network, you can configure DR priorities for interfaces to affect DR or BDR election.

After an interface is disabled from sending OSPFv3 packets, other routers cannot obtain any information from the interface.

Configuration prerequisites

Before you tune and optimize OSPFv3 networks, complete the following tasks:

- Enable IPv6 packet forwarding.
- Configure OSPFv3 basic functions.

Configuring OSPFv3 timers

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the hello interval.	ospfv3 timer hello <i>seconds</i> [instance <i>instance-id</i>]	Optional. By default, the hello interval on P2P, broadcast interfaces is 10 seconds.
4. Specify the poll interval.	ospfv3 timer poll <i>seconds</i> [instance <i>instance-id</i>]	Optional. By default, the poll interval is 120 seconds.
5. Configure the dead interval.	ospfv3 timer dead <i>seconds</i> [instance <i>instance-id</i>]	Optional. By default, the dead interval on P2P, broadcast interfaces is 40 seconds. The dead interval set on neighboring interfaces cannot be too short. Otherwise, a neighbor is easily considered down.
6. Configure the LSA retransmission interval.	ospfv3 timer retransmit <i>interval</i> [instance <i>instance-id</i>]	Optional. By default, the LSA retransmission interval is 5 seconds. The LSA retransmission interval cannot be too short; otherwise, unnecessary retransmissions occur.
7. Configure the LSA transmission delay.	ospfv3 trans-delay <i>seconds</i> [instance <i>instance-id</i>]	Optional. By default, the LSA transmission delay is 1 second.
8. Return to system view.	quit	N/A
9. Enter OSPFv3 view.	ospfv3 [<i>process-id</i>]	N/A
10. Configure the SPF timers.	spf timers <i>delay-interval</i> <i>hold-interval</i>	Optional. By default, <i>delay-interval</i> is 5 seconds and <i>hold-interval</i> is 10 seconds. Setting both the <i>delay-interval</i> and <i>hold-interval</i> to 0 triggers an SPF calculation at once, improving the network convergence speed.

Configuring a DR priority for an interface

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A

Step	Command	Remarks
3. Configure a DR priority.	ospfv3 dr-priority <i>priority</i> [instance <i>instance-id</i>]	Optional. By default, the DR priority is 1. The DR priority of an interface determines the interface's qualification in DR election. Interfaces having the priority 0 cannot become a DR or BDR.

Ignoring MTU check for DD packets

When LSAs are few in DD packets, it is unnecessary to check the MTU in DD packets to improve efficiency.

To ignore MTU check for DD packets:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Ignore MTU check for DD packets.	ospfv3 mtu-ignore [instance <i>instance-id</i>]	Not ignored by default.

Disabling interfaces from receiving and sending OSPFv3 packets

Follow these guidelines when you disable interfaces from receiving and sending OSPF packets:

- Multiple OSPFv3 processes can disable the same interface from receiving and sending OSPFv3 packets. Using the **silent-interface** command disables only the interfaces associated with the current process.
- After an OSPF interface is set to **silent**, direct routes of the interface can still be advertised in Intra-Area-Prefix-LSAs via other interfaces, but other OSPFv3 packets cannot be advertised. No neighboring relationship can be established on the interface. This feature can enhance the adaptability of OSPFv3 networking.

To disable interfaces from receiving and sending OSPFv3 packets:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPFv3 view.	ospfv3 [<i>process-id</i>]	N/A
3. Disable interfaces from receiving and sending OSPFv3 packets.	silent-interface { <i>interface-type</i> <i>interface-number</i> all }	Not disabled by default.

Enabling the logging of neighbor state changes

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPFv3 view.	ospfv3 [<i>process-id</i>]	N/A
3. Enable the logging of neighbor state changes.	log-peer-change	Enabled by default.

Configuring OSPFv3 GR

ⓘ IMPORTANT:

You cannot configure OSPFv3 GR after configuring OSPFv3 virtual links, because they are not supported at the same time.

Graceful Restart ensures the continuity of packet forwarding when a routing protocol restarts or a master/slave switchover occurs:

- **GR Restarter**—Graceful restarting router. It must be Graceful Restart capable.
- **GR Helper**—The neighbor of the GR Restarter. It helps the GR Restarter to complete the GR process.

To prevent service interruption after a master/slave switchover, a GR Restarter running OSPFv3 must complete the following tasks:

- Keep the GR Restarter forwarding entries stable during reboot.
- Establish all adjacencies and obtain complete topology information after reboot.

After the master/slave switchover, the GR Restarter sends a Grace-LSA to tell its neighbors that it performs a GR. Upon receiving the Grace-LSA, the neighbors with the GR Helper capability enter the helper mode (and are then called "GR Helpers"). Then, the GR Restarter retrieves its adjacencies and LSDB with the help of the GR Helpers.

Configuring GR Restarter

You can configure the GR Restarter capability on a GR Restarter.

To configure GR Restarter:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPFv3 view.	ospfv3 [<i>process-id</i>]	N/A
3. Enable the GR capability.	graceful-restart enable	Disabled by default.
4. Configure the GR interval.	graceful-restart interval <i>interval-value</i>	Optional. 120 seconds by default.

Configuring GR Helper

You can configure the GR Helper capability on a GR Helper.

To configure GR Helper

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPFv3 view.	ospfv3 [<i>process-id</i>]	N/A
3. Enable the GR Helper capability.	graceful-restart helper enable	Optional. Enabled by default.
4. Enable strict LSA checking.	graceful-restart helper strict-lsa-checking	Optional. Disabled by default.

Configuring BFD for OSPFv3

Bidirectional forwarding detection (BFD) provides a mechanism to quickly detect the connectivity of links between OSPFv3 neighbors, thus to improve the convergence speed of OSPFv3. For more information about BFD, see *High Availability Configuration Guide*.

After discovering neighbors by sending hello packets, OSPFv3 notifies BFD of the neighbor addresses, and BFD uses these addresses to establish sessions. Before a BFD session is established, it is in the down state. In this state, BFD control packets are sent at an interval of no less than one second to reduce BFD control packet traffic. After the BFD session is established, BFD control packets are sent at the negotiated interval, thereby implementing fast fault detection.

To configure BFD for OSPFv3, you need to configure OSPFv3 first.

To configure BFD for OSPFv3:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPFv3 view.	ospfv3 [<i>process-id</i>]	N/A
3. Specify a router ID.	router-id <i>router-id</i>	N/A
4. Quit the OSPFv3 view.	quit	N/A
5. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
6. Enable an OSPFv3 process on the interface.	ospfv3 <i>process-id</i> area <i>area-id</i> [instance <i>instance-id</i>]	Not enabled by default.
7. Enable BFD on the interface.	ospfv3 bfd enable [instance <i>instance-id</i>]	Not enabled by default.

Applying IPsec policies for OSPFv3

To protect routing information and defend attacks, OSPFv3 can authenticate protocol packets by using an IPsec policy.

Outbound OSPFv3 packets carry the Security Parameter Index (SPI) defined in the relevant IPsec policy. A device uses the SPI carried in a received packet to match against the configured IPsec policy. If they match, the device accepts the packet; otherwise, it discards the packet and will not establish a neighbor relationship with the sending device.

You can configure an IPsec policy for an area, an interface or a virtual link.

- To implement area-based IPsec protection, you need to configure the same IPsec policy on the routers in the target area.
- To implement interface-based IPsec protection, you need to configure the same IPsec policy on the interfaces between two neighboring routers.
- To implement virtual link-based IPsec protection, you need to configure the same IPsec policy on the two routers connected over the virtual link.

If an interface and its area each have an IPsec policy configured, the interface uses its own IPsec policy. If a virtual link and area 0 each have an IPsec policy configured, the virtual link uses its own IPsec policy.

Configuration prerequisites

Before you apply an IPsec policy for OSPFv3, complete following tasks.

- Create an IPsec proposal.
- Create an IPsec policy.

For more information about IPsec policy configuration, see *Security Configuration Guide*.

Configuration guidelines

An IPsec policy used for OSPFv3 can only be in manual mode. For more information, see *Security Configuration Guide*.

Configuration procedure

To apply an IPsec policy in an area:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPFv3 view.	ospfv3 [<i>process-id</i>]	N/A
3. Enter OSPF area view.	area <i>area-id</i>	N/A
4. Apply an IPsec policy in the area.	enable ipsec-policy <i>policy-name</i>	Not configured by default.

To apply an IPsec policy on an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Apply an IPsec policy on the interface.	ospfv3 ipsec-policy <i>policy-name</i> [instance <i>instance-id</i>]	Not configured by default

To apply an IPsec policy on a virtual link:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter OSPFv3 view.	ospfv3 [<i>process-id</i>]	N/A
3. Enter OSPF area view.	area <i>area-id</i>	N/A

Step	Command	Remarks
4. Apply an IPsec policy on a virtual link.	vlink-peer <i>router-id</i> [hello <i>seconds</i> retransmit <i>seconds</i> trans-delay <i>seconds</i> dead <i>seconds</i> instance <i>instance-id</i> ipsec-policy <i>policy-name</i>] *	Not configured by default.

Displaying and maintaining OSPFv3

Task	Command	Remarks
Display OSPFv3 process brief information.	display ospfv3 [<i>process-id</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPFv3 interface information.	display ospfv3 interface [<i>interface-type</i> <i>interface-number</i> statistic] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPFv3 LSDB information.	display ospfv3 [<i>process-id</i>] lsdb [[external inter-prefix inter-router intra-prefix link network router grace] [<i>link-state-id</i>] [originate-router <i>router-id</i>] total] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPFv3 LSDB statistics.	display ospfv3 lsdb statistic [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPFv3 neighbor information.	display ospfv3 [<i>process-id</i>] [area <i>area-id</i>] peer [[<i>interface-type</i> <i>interface-number</i>] [verbose] <i>peer-router-id</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPFv3 neighbor statistics.	display ospfv3 peer statistic [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPFv3 routing table information.	display ospfv3 [<i>process-id</i>] routing [<i>ipv6-address</i> <i>prefix-length</i> <i>ipv6-address/prefix-length</i> abr-routes asbr-routes all statistics] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPFv3 area topology information.	display ospfv3 [<i>process-id</i>] topology [area <i>area-id</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPFv3 virtual link information.	display ospfv3 [<i>process-id</i>] vlink [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPFv3 next hop information.	display ospfv3 [<i>process-id</i>] next-hop [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPFv3 link state request list information.	display ospfv3 [<i>process-id</i>] request-list [{ external inter-prefix inter-router intra-prefix link network router grace] [<i>link-state-id</i>] [originate-router <i>ip-address</i>] statistics] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display OSPFv3 link state retransmission list information.	display ospfv3 [<i>process-id</i>] retrans-list [{ external inter-prefix inter-router intra-prefix link network router grace] [<i>link-state-id</i>] [originate-router <i>ip-address</i>] statistics] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display OSPFv3 statistics.	display ospfv3 statistic [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the GR status of the specified OSPFv3 process.	display ospfv3 [<i>process-id</i>] graceful-restart status [{ begin exclude include } <i>regular-expression</i>]	Available in any view

OSPFv3 configuration examples

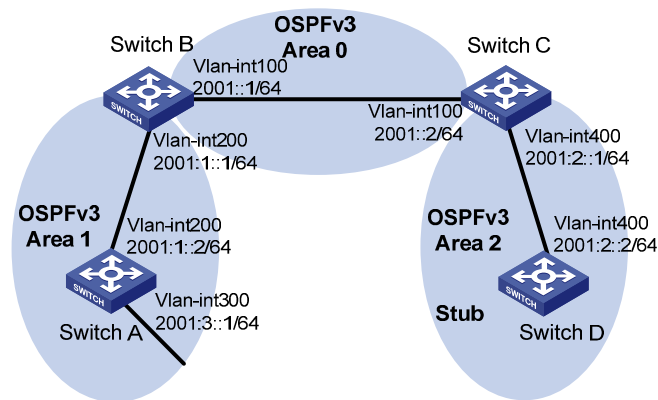
Configuring OSPFv3 areas

Network requirements

In Figure 107, all switches run OSPFv3. The AS is split into three areas, in which, Switch B and Switch C act as ABRs to forward routing information between areas.

Configure Area 2 as a stub area in order to reduce LSAs in the area without affecting route reachability.

Figure 107 Network diagram



Configuration procedure

1. Configure IPv6 addresses for interfaces. (Details not shown.)
2. Configure OSPFv3 basic functions:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] ospfv3
[SwitchA-ospfv3-1] router-id 1.1.1.1
[SwitchA-ospfv3-1] quit
[SwitchA] interface vlan-interface 300
[SwitchA-Vlan-interface300] ospfv3 1 area 1
[SwitchA-Vlan-interface300] quit
[SwitchA] interface vlan-interface 200
[SwitchA-Vlan-interface200] ospfv3 1 area 1
[SwitchA-Vlan-interface200] quit
```

Configure Switch B.

```
<SwitchB> system-view
```

```
[SwitchB] ipv6
[SwitchB] ospfv3
[SwitchB-ospf-1] router-id 2.2.2.2
[SwitchB-ospf-1] quit
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ospfv3 1 area 0
[SwitchB-Vlan-interface100] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] ospfv3 1 area 1
[SwitchB-Vlan-interface200] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] ospfv3
[SwitchC-ospfv3-1] router-id 3.3.3.3
[SwitchC-ospfv3-1] quit
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] ospfv3 1 area 0
[SwitchC-Vlan-interface100] quit
[SwitchC] interface vlan-interface 400
[SwitchC-Vlan-interface400] ospfv3 1 area 2
[SwitchC-Vlan-interface400] quit
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] ipv6
[SwitchD] ospfv3
[SwitchD-ospfv3-1] router-id 4.4.4.4
[SwitchD-ospfv3-1] quit
[SwitchD] interface Vlan-interface 400
[SwitchD-Vlan-interface400] ospfv3 1 area 2
[SwitchD-Vlan-interface400] quit
```

Display OSPFv3 neighbor information on Switch B.

```
[SwitchB] display ospfv3 peer
```

```

                OSPFv3 Area ID 0.0.0.0 (Process 1)
-----
Neighbor ID    Pri   State           Dead Time   Interface   Instance ID
3.3.3.3        1     Full/DR         00:00:39   Vlan100    0
```

```

                OSPFv3 Area ID 0.0.0.1 (Process 1)
-----
Neighbor ID    Pri   State           Dead Time   Interface   Instance ID
1.1.1.1        1     Full/Backup     00:00:38   Vlan200    0
```

Display OSPFv3 neighbor information on Switch C.

```
[SwitchC] display ospfv3 peer
```

```

                OSPFv3 Area ID 0.0.0.0 (Process 1)
-----
Neighbor ID    Pri   State           Dead Time   Interface   Instance ID
```



```
2.2.2.2          1      Full/Backup    00:00:39    Vlan100      0
```

```
OSPFv3 Area ID 0.0.0.2 (Process 1)
```

```
-----  
Neighbor ID      Pri   State           Dead Time   Interface   Instance ID  
4.4.4.4          1     Full/DR         00:00:38   Vlan400     0
```

```
# Display OSPFv3 routing table information on Switch D.
```

```
[SwitchD] display ospfv3 routing
```

```
E1 - Type 1 external route,   IA - Inter area route,   I - Intra area route  
E2 - Type 2 external route,   * - Selected route
```

```
OSPFv3 Router with ID (4.4.4.4) (Process 1)
```

```
-----  
*Destination: 2001::/64  
Type           : IA                               Cost       : 2  
NextHop        : FE80::F40D:0:93D0:1             Interface: Vlan400  
  
*Destination: 2001:1::/64  
Type           : IA                               Cost       : 3  
NextHop        : FE80::F40D:0:93D0:1             Interface: Vlan400  
  
*Destination: 2001:2::/64  
Type           : I                               Cost       : 1  
NextHop        : directly-connected              Interface: Vlan400  
  
*Destination: 2001:3::/64  
Type           : IA                               Cost       : 4  
NextHop        : FE80::F40D:0:93D0:1             Interface: Vlan400
```

3. Configure Area 2 as a stub area:

```
# Configure Switch D
```

```
[SwitchD] ospfv3
```

```
[SwitchD-ospfv3-1] area 2
```

```
[SwitchD-ospfv3-1-area-0.0.0.2] stub
```

```
# Configure Switch C, and specify the cost of the default route sent to the stub area as 10.
```

```
[SwitchC] ospfv3
```

```
[SwitchC-ospfv3-1] area 2
```

```
[SwitchC-ospfv3-1-area-0.0.0.2] stub
```

```
[SwitchC-ospfv3-1-area-0.0.0.2] default-cost 10
```

```
# Display OSPFv3 routing table information on Switch D. A default route is added, and its cost is the cost of a direct route plus the configured cost.
```

```
[SwitchD] display ospfv3 routing
```

```
E1 - Type 1 external route,   IA - Inter area route,   I - Intra area route  
E2 - Type 2 external route,   * - Selected route
```

```
OSPFv3 Router with ID (4.4.4.4) (Process 1)
```

```

*Destination: ::/0
  Type       : IA                               Cost       : 11
  NextHop    : FE80::F40D:0:93D0:1           Interface:  Vlan400

*Destination: 2001::/64
  Type       : IA                               Cost       : 2
  NextHop    : FE80::F40D:0:93D0:1           Interface:  Vlan400

*Destination: 2001:1::/64
  Type       : IA                               Cost       : 3
  NextHop    : FE80::F40D:0:93D0:1           Interface:  Vlan400

*Destination: 2001:2::/64
  Type       : I                                Cost       : 1
  NextHop    : directly-connected            Interface:  Vlan400

*Destination: 2001:3::/64
  Type       : IA                               Cost       : 4
  NextHop    : FE80::F40D:0:93D0:1           Interface:  Vlan400

```

4. Configure Area 2 as a totally stub area:

Configure Area 2 as a totally stub area on Switch C.

```
[SwitchC-ospfv3-1-area-0.0.0.2] stub no-summary
```

Display OSPFv3 routing table information on Switch D. Route entries are reduced. All non-direct routes are removed except the default route.

```
[SwitchD] display ospfv3 routing
```

```

E1 - Type 1 external route,   IA - Inter area route,   I - Intra area route
E2 - Type 2 external route,   * - Selected route

```

```
OSPFv3 Router with ID (4.4.4.4) (Process 1)
```

```

-----
*Destination: ::/0
  Type       : IA                               Cost       : 11
  NextHop    : FE80::F40D:0:93D0:1           Interface:  Vlan400

*Destination: 2001:2::/64
  Type       : I                                Cost       : 1
  NextHop    : directly-connected            Interface:  Vlan400

```

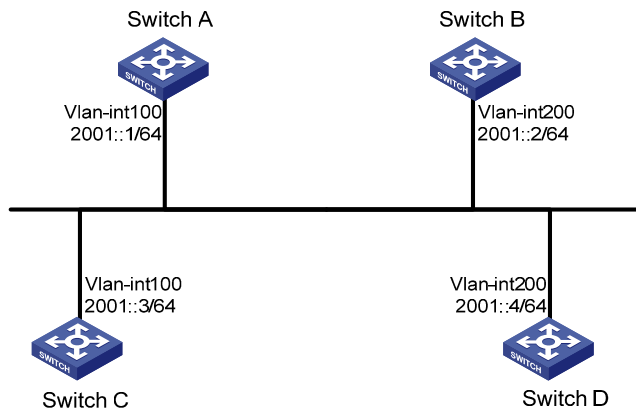
Configuring OSPFv3 DR election

Network requirements

As shown in [Figure 108](#):

- The priority of Switch A is 100, the highest priority on the network, so it will be the DR.
- The priority of Switch C is 2, the second highest priority on the network, so it will be the BDR.
- The priority of Switch B is 0, so it cannot become the DR.
- Router D has the default priority 1.

Figure 108 Network diagram



Configuration procedure

1. Configure IPv6 addresses for interfaces. (Details not shown.)
2. Configure OSPFv3 basic functions:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] ospfv3
[SwitchA-ospfv3-1] router-id 1.1.1.1
[SwitchA-ospfv3-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ospfv3 1 area 0
[SwitchA-Vlan-interface100] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] ospfv3
[SwitchB-ospfv3-1] router-id 2.2.2.2
[SwitchB-ospfv3-1] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] ospfv3 1 area 0
[SwitchB-Vlan-interface200] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] ospfv3
[SwitchC-ospfv3-1] router-id 3.3.3.3
[SwitchC-ospfv3-1] quit
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] ospfv3 1 area 0
[SwitchC-Vlan-interface100] quit
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] ipv6
```

```
[SwitchD] ospfv3
[SwitchD-ospfv3-1] router-id 4.4.4.4
[SwitchD-ospfv3-1] quit
[SwitchD] interface vlan-interface 200
[SwitchD-Vlan-interface200] ospfv3 1 area 0
[SwitchD-Vlan-interface200] quit
```

Display neighbor information on Switch A. The switches have the same default DR priority 1. Then, Switch D (the switch with the highest Router ID) is elected as the DR, and Switch C is the BDR.

```
[SwitchA] display ospfv3 peer
      OSPFv3 Area ID 0.0.0.0 (Process 1)
```

```
-----
Neighbor ID      Pri   State           Dead Time   Interface   Instance ID
2.2.2.2          1     2-Way/DROther   00:00:36   Vlan200     0
3.3.3.3          1     Full/Backup     00:00:35   Vlan100     0
4.4.4.4          1     Full/DR         00:00:33   Vlan200     0
```

Display neighbor information on Switch D. The neighbor states are all full.

```
[SwitchD] display ospfv3 peer
      OSPFv3 Area ID 0.0.0.0 (Process 1)
```

```
-----
Neighbor ID      Pri   State           Dead Time   Interface   Instance ID
1.1.1.1          1     Full/DROther    00:00:30   Vlan100     0
2.2.2.2          1     Full/DROther    00:00:37   Vlan200     0
3.3.3.3          1     Full/Backup     00:00:31   Vlan100     0
```

3. Configure DR priorities for interfaces:

Configure the DR priority of VLAN-interface 100 as 100 on Switch A.

```
[SwitchA] interface Vlan-interface 100
[SwitchA-Vlan-interface100] ospfv3 dr-priority 100
[SwitchA-Vlan-interface100] quit
```

Configure the DR priority of VLAN-interface 200 as 0 on Switch B.

```
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] ospfv3 dr-priority 0
[SwitchB-Vlan-interface200] quit
```

Configure the DR priority of VLAN-interface 100 of Switch C as 2.

```
[SwitchC] interface Vlan-interface 100
[SwitchC-Vlan-interface100] ospfv3 dr-priority 2
[SwitchC-Vlan-interface100] quit
```

Display neighbor information on Switch A. DR priorities have been updated, but the DR and BDR are not changed.

```
[SwitchA] display ospfv3 peer
      OSPFv3 Area ID 0.0.0.0 (Process 1)
```

```
-----
Neighbor ID      Pri   State           Dead Time   Interface   Instance ID
2.2.2.2          0     2-Way/DROther   00:00:38   Vlan200     0
3.3.3.3          2     Full/Backup     00:00:32   Vlan100     0
4.4.4.4          1     Full/DR         00:00:36   Vlan200     0
```

Display neighbor information on Switch D. Switch D is still the DR.

```
[SwitchD] display ospfv3 peer
```

```

OSPFv3 Area ID 0.0.0.0 (Process 1)
-----
Neighbor ID      Pri   State           Dead Time   Interface   Instance ID
1.1.1.1          100  Full/DROther    00:00:33   Vlan100    0
2.2.2.2          0     Full/DROther    00:00:36   Vlan200    0
3.3.3.3          2     Full/Backup     00:00:40   Vlan100    0

```

4. Restart DR and BDR election:

Use the **shutdown** and **undo shutdown** commands on interfaces to restart DR and BDR election. (Details not shown.)

Display neighbor information on Switch A. Switch C becomes the BDR.

```

[SwitchA] display ospfv3 peer
OSPFv3 Area ID 0.0.0.0 (Process 1)
-----
Neighbor ID      Pri   State           Dead Time   Interface   Instance ID
2.2.2.2          0     Full/DROther    00:00:31   Vlan200    0
3.3.3.3          2     Full/Backup     00:00:39   Vlan100    0
4.4.4.4          1     Full/DROther    00:00:37   Vlan200    0

```

Display neighbor information on Switch D. Switch A becomes the DR.

```

[SwitchD] display ospfv3 peer
OSPFv3 Area ID 0.0.0.0 (Process 1)
-----
Neighbor ID      Pri   State           Dead Time   Interface   Instance ID
1.1.1.1          100  Full/DR         00:00:34   Vlan100    0
2.2.2.2          0     2-Way/DROther  00:00:34   Vlan200    0
3.3.3.3          2     Full/Backup     00:00:32   Vlan100    0

```

Configuring OSPFv3 route redistribution

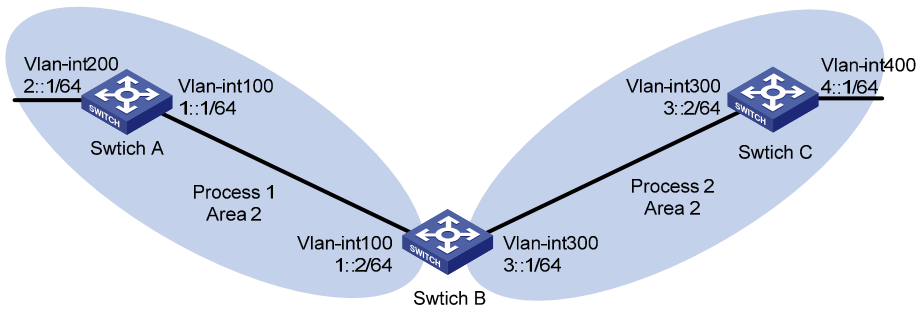
Network requirements

As shown in [Figure 109](#):

- Switch A, Switch B, and Switch C are in Area 2.
- OSPFv3 process 1 and OSPFv3 process 2 are enabled on Switch B. Switch B communicates with Switch A and Switch C through OSPFv3 process 1 and OSPFv3 process 2.

Configure OSPFv3 process 2 to redistribute direct routes and the routes from OSPFv3 process 1 on Switch B and set the default metric for redistributed routes to 3. Then, Switch C can learn the routes destined for 1::0/64 and 2::0/64, and Switch A cannot learn the routes destined for 3::0/64 or 4::0/64.

Figure 109 Network diagram



Configuration procedure

1. Configure IPv6 addresses for interfaces. (Details not shown.)
2. Configure OSPFv3 basic functions:

Enable OSPFv3 process 1 on Switch A.

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] ospfv3 1
[SwitchA-ospfv3-1] router-id 1.1.1.1
[SwitchA-ospfv3-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ospfv3 1 area 2
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 200
[SwitchA-Vlan-interface200] ospfv3 1 area 2
[SwitchA-Vlan-interface200] quit
```

Enable OSPFv3 process 1 and OSPFv3 process 2 on Switch B.

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] ospfv3 1
[SwitchB-ospfv3-1] router-id 2.2.2.2
[SwitchB-ospfv3-1] quit
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ospfv3 1 area 2
[SwitchB-Vlan-interface100] quit
[SwitchB] ospfv3 2
[SwitchB-ospfv3-2] router-id 3.3.3.3
[SwitchB-ospfv3-2] quit
[SwitchB] interface vlan-interface 300
[SwitchB-Vlan-interface300] ospfv3 2 area 2
[SwitchB-Vlan-interface300] quit
```

Enable OSPFv3 process 2 on Switch C.

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] ospfv3 2
[SwitchC-ospfv3-2] router-id 4.4.4.4
[SwitchC-ospfv3-2] quit
```

```
[SwitchC] interface vlan-interface 300
[SwitchC-Vlan-interface300] ospfv3 2 area 2
[SwitchC-Vlan-interface300] quit
[SwitchC] interface vlan-interface 400
[SwitchC-Vlan-interface400] ospfv3 2 area 2
[SwitchC-Vlan-interface400] quit
```

Display the routing table of Switch C.

```
[SwitchC] display ipv6 routing-table
```

Routing Table :

Destinations : 6 Routes : 6

```
Destination: ::1/128                      Protocol : Direct
NextHop      : ::1                        Preference: 0
Interface    : InLoop0                   Cost      : 0
```

```
Destination: 3::/64                       Protocol : Direct
NextHop      : 3::2                       Preference: 0
Interface    : Vlan300                    Cost      : 0
```

```
Destination: 3::2/128                    Protocol : Direct
NextHop      : ::1                        Preference: 0
Interface    : InLoop0                   Cost      : 0
```

```
Destination: 4::/64                       Protocol : Direct
NextHop      : 4::1                       Preference: 0
Interface    : Vlan400                    Cost      : 0
```

```
Destination: 4::1/128                    Protocol : Direct
NextHop      : ::1                        Preference: 0
Interface    : InLoop0                   Cost      : 0
```

```
Destination: FE80::/10                   Protocol : Direct
NextHop      : ::                         Preference: 0
Interface    : NULL0                      Cost      : 0
```

3. Configure OSPFv3 route redistribution:

Configure OSPFv3 process 2 to redistribute direct routes and the routes from OSPFv3 process 1 on Switch B.

```
[SwitchB] ospfv3 2
[SwitchB-ospfv3-2] default cost 3
[SwitchB-ospfv3-2] import-route ospfv3 1
[SwitchB-ospfv3-2] import-route direct
[SwitchB-ospfv3-2] quit
```

Display the routing table of Switch C.

```
[SwitchC] display ipv6 routing-table
```

Routing Table :

Destinations : 8 Routes : 8

```
Destination: ::1/128                      Protocol : Direct
```

```

NextHop      : ::1                               Preference: 0
Interface    : InLoop0                          Cost       : 0

Destination: 1::/64                             Protocol   : OSPFv3
NextHop      : FE80::200:CFF:FE01:1C03          Preference: 150
Interface    : Vlan300                          Cost       : 3

Destination: 2::/64                             Protocol   : OSPFv3
NextHop      : FE80::200:CFF:FE01:1C03          Preference: 150
Interface    : Vlan300                          Cost       : 3

Destination: 3::/64                             Protocol   : Direct
NextHop      : 3::2                             Preference: 0
Interface    : Vlan300                          Cost       : 0

Destination: 3::2/128                           Protocol   : Direct
NextHop      : ::1                               Preference: 0
Interface    : InLoop0                          Cost       : 0

Destination: 4::/64                             Protocol   : Direct
NextHop      : 4::1                             Preference: 0
Interface    : Vlan400                          Cost       : 0

Destination: 4::1/128                           Protocol   : Direct
NextHop      : ::1                               Preference: 0
Interface    : InLoop0                          Cost       : 0

Destination: FE80::/10                          Protocol   : Direct
NextHop      : ::                               Preference: 0
Interface    : NULL0                             Cost       : 0

```

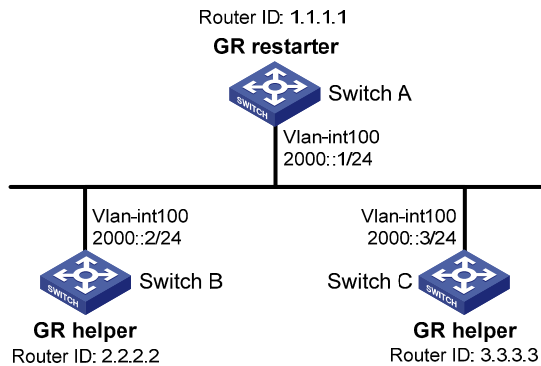
Configuring OSPFv3 GR

Network requirements

As shown in [Figure 110](#), Switch A, Switch B, and Switch C that belong to the same AS and the same OSPFv3 routing domain are GR capable.

Switch A acts as the GR Restarter. Switch B and Switch C are the GR Helpers and synchronize their LSDBs with Switch A through out-of-band (OOB) communication of GR.

Figure 110 Network diagram



Configuration procedure

1. Configure IPv6 addresses for interfaces. (Details not shown.)
2. Configure OSPFv3 basic functions:
 - # On Switch A, enable OSPFv3 process 1, enable GR, and set the router ID to 1.1.1.1.

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] ospfv3 1
[SwitchA-ospfv3-1] router-id 1.1.1.1
[SwitchA-ospfv3-1] graceful-restart enable
[SwitchA-ospfv3-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ospfv3 1 area 1
[SwitchA-Vlan-interface100] quit
```
 - # Enable OSPFv3 on Switch B and set the router ID to 2.2.2.2. (By default, GR helper is enabled on Switch B).

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] ospfv3 1
[SwitchB-ospfv3-1] router-id 2.2.2.2
[SwitchB-ospfv3-1] quit
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ospfv3 1 area 1
[SwitchB-Vlan-interface100] quit
```
 - # Enable OSPFv3 on Switch C and set the router ID to 3.3.3.3. (By default, GR helper is enabled on Switch C).

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] ospfv3 1
[SwitchC-ospfv3-1] router-id 3.3.3.3
[SwitchC-ospfv3-1] quit
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] ospfv3 1 area 1
[SwitchC-Vlan-interface100] quit
```
3. Verify the configuration:

After all switches function properly, perform a master/slave switchover on Switch A to trigger an OSPFv3 GR operation.

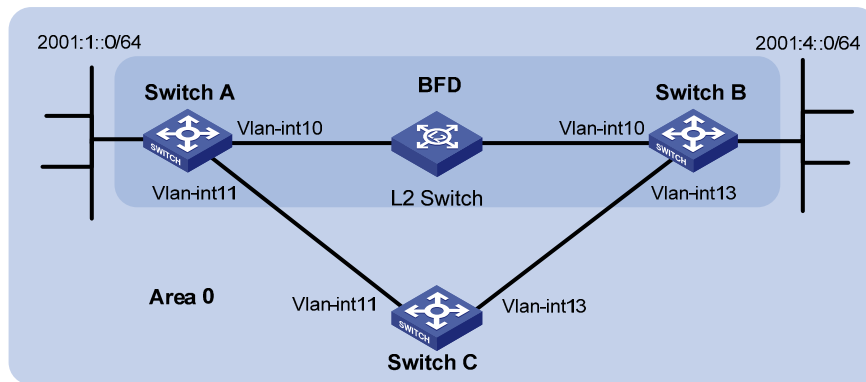
Configuring BFD for OSPFv3

Network requirements

In Figure 111, configure OSPFv3 on Switch A, Switch B and Switch C and configure BFD over the link Switch A ↔ L2 Switch ↔ Switch B.

After the link Switch A ↔ L2 Switch ↔ Switch B fails, BFD can quickly detect the failure and notify OSPFv3 of the failure. Then Switch A and Switch B communicate through Switch C.

Figure 111 Network diagram



Device	Interface	IPv6 address	Device	Interface	IPv6 address
Switch A	Vlan-int10	2001::1/64	Switch B	Vlan-int10	2001::2/64
	Vlan-int11	2001:2::1/64		Vlan-int13	2001:3::2/64
Switch C	Vlan-int11	2001:2::2/64			
	Vlan-int13	2001:3::1/64			

Configuration procedure

1. Configure IP addresses for the interfaces. (Details not shown.)
2. Configure OSPF basic functions:


```
# Configure Switch A. Enable OSPFv3 and configure the router ID as 1.1.1.1.
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] ospfv3
[SwitchA-ospfv3-1] router-id 1.1.1.1
[SwitchA-ospfv3-1] quit
[SwitchA] interface vlan-interface 10
[SwitchA-Vlan-interface10] ospfv3 1 area 0
[SwitchA-Vlan-interface10] quit
[SwitchA] interface vlan-interface 11
[SwitchA-Vlan-interface11] ospfv3 1 area 0
[SwitchA-Vlan-interface11] quit

# Configure Switch B. Enable OSPFv3 and configure the router ID as 2.2.2.2.
<SwitchB> system-view
[SwitchB] ipv6
```

```

[SwitchB] ospfv3
[SwitchB-ospf-1] router-id 2.2.2.2
[SwitchB-ospf-1] quit
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] ospfv3 1 area 0
[SwitchB-Vlan-interface10] quit
[SwitchB] interface vlan-interface 13
[SwitchB-Vlan-interface13] ospfv3 1 area 0
[SwitchB-Vlan-interface13] quit

```

Configure Switch C. Enable OSPFv3 and configure the router ID as 3.3.3.3.

```

<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] ospfv3
[SwitchC-ospfv3-1] router-id 3.3.3.3
[SwitchC-ospfv3-1] quit
[SwitchC] interface vlan-interface 11
[SwitchC-Vlan-interface11] ospfv3 1 area 0
[SwitchC-Vlan-interface11] quit
[SwitchC] interface vlan-interface 13
[SwitchC-Vlan-interface13] ospfv3 1 area 0
[SwitchC-Vlan-interface13] quit

```

3. Configure BFD:

Enable BFD on Switch A and configure BFD parameters.

```

[SwitchA] bfd session init-mode active
[SwitchA] interface vlan-interface 10
[SwitchA-Vlan-interface10] ospfv3 bfd enable
[SwitchA-Vlan-interface10] bfd min-transmit-interval 500
[SwitchA-Vlan-interface10] bfd min-receive-interval 500
[SwitchA-Vlan-interface10] bfd detect-multiplier 7
[SwitchA-Vlan-interface10] return

```

Enable BFD on Switch B and configure BFD parameters.

```

[SwitchB] bfd session init-mode active
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] ospfv3 bfd enable
[SwitchB-Vlan-interface10] bfd min-transmit-interval 500
[SwitchB-Vlan-interface10] bfd min-receive-interval 500
[SwitchB-Vlan-interface10] bfd detect-multiplier 6

```

4. Verify the configuration:

The following operations are performed on Switch A. The operations on Switch B are similar, and are not shown.

Display the BFD information of Switch A.

```

<SwitchA> display bfd session
Total Session Num: 1          Init Mode: Active

IPv6 Session Working Under Ctrl Mode:

Local Discr: 1441          Remote Discr: 1450

```

Source IP: FE80::20F:FF:FE00:1202 (link-local address of VLAN-interface 10 on Switch A)

Destination IP: FE80::20F:FF:FE00:1200 (link-local address of VLAN-interface 10 on Switch B)

Session State: Up Interface: Vlan10
Hold Time: /

Display routes to 2001:4::0/64 on Switch A, and you can see that Switch A communicates with Switch B through the Layer 2 switch.

```
<SwitchA> display ipv6 routing-table 2001:4::0 64 verbose
```

Routing Table :

Summary Count : 2

```
Destination : 2001:4:: PrefixLength : 64
NextHop      : 2001::2 Preference : 10
IpPrecedence : QoS LcId :
RelayNextHop : :: Tag : 0H
Neighbor     : :: ProcessID : 0
Interface    : Vlan-interface10 Protocol : OSPFv3
State        : Active Adv Cost : 1
Tunnel ID    : 0x0 Label : NULL
Age          : 4538sec
```

```
Destination : 2001:4:: PrefixLength : 64
NextHop      : 2001:2::2 Preference : 10
IpPrecedence : QoS LcId :
RelayNextHop : :: Tag : 0H
Neighbor     : :: ProcessID : 0
Interface    : Vlan-interface11 Protocol : OSPFv3
State        : Invalid Adv Cost : 2
Tunnel ID    : 0x0 Label : NULL
Age          : 4515sec
```

Enable BFD debugging on Switch A.

```
<SwitchA> debugging bfd scm
```

```
<SwitchA> debugging bfd event
```

```
<SwitchA> debugging ospfv3 event bfd
```

```
<SwitchA> terminal debugging
```

After the link between Switch B and the Layer 2 switch fails, Switch A quickly detects the change on Switch B.

```
%Nov 5 11:37:43:062 2009 SwitchA BFD/5/BFD_CHANGE_FSM: Sess[FE80::20F:FF:FE00:1202/FE80::20F:FF:FE00:1200,15/15,Vlan10,Ctrl], Sta: UP->DOWN, Diag: 1
```

```
%Nov 5 11:37:43:062 2009 SwitchA OSPFV3/5/OSPFv3_NBR_CHG: OSPFv3 1 Neighbor 2.2.2.2(Vlan-interface10) from Full to Down.
```

```
*Nov 5 11:37:43:062 2009 SwitchA RM/6/RMDEBUG: OSPFv3 OSPFv3-BFD: Message Type rcv BFD down, Connect Type direct-connect, Src IP Address FE80::20F:FF:FE00:1202, Dst IP Address FE80::20F:FF:FE00:1200.
```

```
*Nov 5 11:37:43:062 2009 SwitchA RM/6/RMDEBUG: OSPFv3 OSPFv3-BFD: Message Type delete session, Connect Type direct-connect, Src IP Address FE80::20F:FF:FE00:1202, Dst IP Address FE80::20F:FF:FE00:1200.
```

Display the BFD information of Switch A. You can see that Switch A has removed its neighbor relationship with Switch B and therefore no information is output.

```
<SwitchA> display bfd session
```

Display routes to 2001:4::0/64 on Switch A, and you can see that Switch A communicates with Switch B through Switch C.

```
<SwitchA> display ipv6 routing-table 2001:4::0 64 verbose
```

```
Routing Table :
```

```
Summary Count : 1
```

Destination	: 2001:4::	PrefixLength	: 64
NextHop	: 2001:2::2	Preference	: 10
IpPrecedence	:	QosLcId	:
RelayNextHop	: ::	Tag	: 0H
Neighbor	: ::	ProcessID	: 0
Interface	: Vlan-interface11	Protocol	: OSPFv3
State	: Invalid Adv	Cost	: 2
Tunnel ID	: 0x0	Label	: NULL
Age	: 4610sec		

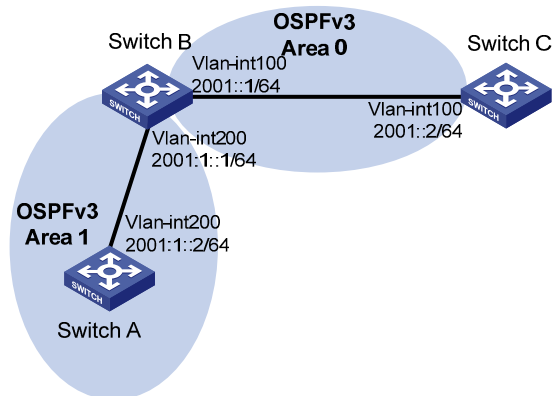
Configuring OSPFv3 IPsec policies

Network requirements

As shown in [Figure 112](#):

- Configure OSPFv3 on the switches. The AS is divided into two areas.
- Configure IPsec policies on the switches to authenticate and encrypt protocol packets.

Figure 112 Network diagram



Configuration procedure

1. Configure IPv6 addresses for interfaces. (Details not shown.)
2. Configure OSPFv3 basic functions:

```
# Configure Switch A: enable OSPFv3 and configure the router ID as 1.1.1.1.
```

```
<SwitchA> system-view
```

```
[SwitchA] ipv6
```

```
[SwitchA] ospfv3 1
```

```
[SwitchA-ospfv3-1] router-id 1.1.1.1
```

```
[SwitchA-ospfv3-1] quit
[SwitchA] interface vlan-interface 200
[SwitchA-Vlan-interface200] ospfv3 1 area 1
[SwitchA-Vlan-interface200] quit
```

Configure Switch B: enable OSPFv3 and configure the router ID as 2.2.2.2.

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] ospfv3 1
[SwitchB-ospfv3-1] router-id 2.2.2.2
[SwitchB-ospfv3-1] quit
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ospfv3 1 area 0
[SwitchB-Vlan-interface100] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] ospfv3 1 area 1
[SwitchB-Vlan-interface200] quit
```

Configure Switch C: enable OSPFv3 and configure the router ID as 3.3.3.3.

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] ospfv3 1
[SwitchC-ospfv3-1] router-id 3.3.3.3
[SwitchC-ospfv3-1] quit
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] ospfv3 1 area 0
[SwitchC-Vlan-interface100] quit
```

3. Configure OSPFv3 IPsec policies:

On Switch A, create an IPsec proposal named **tran1**, and set the encapsulation mode to transport mode, the security protocol to ESP, the encryption algorithm to DES, and authentication algorithm to SHA1; create an IPsec policy named **policy001**, specify the manual mode for it, reference IPsec proposal **tran1**, set the SPIs of the inbound and outbound SAs to 12345, and the keys for the inbound and outbound SAs using ESP to abcdefg.

```
[SwitchA] ipsec proposal tran1
[SwitchA-ipsec-proposal-tran1] encapsulation-mode transport
[SwitchA-ipsec-proposal-tran1] transform esp
[SwitchA-ipsec-proposal-tran1] esp encryption-algorithm des
[SwitchA-ipsec-proposal-tran1] esp authentication-algorithm sha1
[SwitchA-ipsec-proposal-tran1] quit
[SwitchA] ipsec policy policy001 10 manual
[SwitchA-ipsec-policy-manual-policy001-10] proposal tran1
[SwitchA-ipsec-policy-manual-policy001-10] sa spi outbound esp 12345
[SwitchA-ipsec-policy-manual-policy001-10] sa spi inbound esp 12345
[SwitchA-ipsec-policy-manual-policy001-10] sa string-key outbound esp abcdefg
[SwitchA-ipsec-policy-manual-policy001-10] sa string-key inbound esp abcdefg
[SwitchA-ipsec-policy-manual-policy001-10] quit
```

On Switch B, create an IPsec proposal named **tran1**, and set the encapsulation mode to transport mode, the security protocol to ESP, the encryption algorithm to DES, and authentication algorithm to SHA1; create an IPsec policy named **policy001**, specify the manual mode for it, reference IPsec proposal **tran1**, set the SPIs of the inbound and outbound SAs to 12345, and the

keys for the inbound and outbound SAs using ESP to abcdefg; create an IPsec proposal named **tran2**, and set the encapsulation mode to transport mode, the security protocol to ESP, the encryption algorithm to DES, and authentication algorithm to SHA1; create an IPsec policy named **policy002**, specify the manual mode for it, reference IPsec proposal **tran2**, set the SPIs of the inbound and outbound SAs to 54321, and the keys for the inbound and outbound SAs using ESP to gfedcba.

```
[SwitchB] ipsec proposal tran1
[SwitchB-ipsec-proposal-tran1] encapsulation-mode transport
[SwitchB-ipsec-proposal-tran1] transform esp
[SwitchB-ipsec-proposal-tran1] esp encryption-algorithm des
[SwitchB-ipsec-proposal-tran1] esp authentication-algorithm sha1
[SwitchB-ipsec-proposal-tran1] quit
[SwitchB] ipsec policy policy001 10 manual
[SwitchB-ipsec-policy-manual-policy001-10] proposal tran1
[SwitchB-ipsec-policy-manual-policy001-10] sa spi outbound esp 12345
[SwitchB-ipsec-policy-manual-policy001-10] sa spi inbound esp 12345
[SwitchB-ipsec-policy-manual-policy001-10] sa string-key outbound esp abcdefg
[SwitchB-ipsec-policy-manual-policy001-10] sa string-key inbound esp abcdefg
[SwitchB-ipsec-policy-manual-policy001-10] quit
[SwitchB] ipsec proposal tran2
[SwitchB-ipsec-proposal-tran2] encapsulation-mode transport
[SwitchB-ipsec-proposal-tran2] transform esp
[SwitchB-ipsec-proposal-tran2] esp encryption-algorithm des
[SwitchB-ipsec-proposal-tran2] esp authentication-algorithm sha1
[SwitchB-ipsec-proposal-tran2] quit
[SwitchB] ipsec policy policy002 10 manual
[SwitchB-ipsec-policy-manual-policy002-10] proposal tran2
[SwitchB-ipsec-policy-manual-policy002-10] sa spi outbound esp 54321
[SwitchB-ipsec-policy-manual-policy002-10] sa spi inbound esp 54321
[SwitchB-ipsec-policy-manual-policy002-10] sa string-key outbound esp gfedcba
[SwitchB-ipsec-policy-manual-policy002-10] sa string-key inbound esp gfedcba
[SwitchB-ipsec-policy-manual-policy002-10] quit
```

On Switch C, create an IPsec proposal named **tran2**, and set the encapsulation mode to transport mode, the security protocol to ESP, the encryption algorithm to DES, and authentication algorithm to SHA1; create an IPsec policy named **policy002**, specify the manual mode for it, reference IPsec proposal **tran2**, set the SPIs of the inbound and outbound SAs to 54321, and the keys for the inbound and outbound SAs using ESP to gfedcba.

```
[SwitchC] ipsec proposal tran2
[SwitchC-ipsec-proposal-tran2] encapsulation-mode transport
[SwitchC-ipsec-proposal-tran2] transform esp
[SwitchC-ipsec-proposal-tran2] esp encryption-algorithm des
[SwitchC-ipsec-proposal-tran2] esp authentication-algorithm sha1
[SwitchC-ipsec-proposal-tran2] quit
[SwitchC] ipsec policy policy002 10 manual
[SwitchC-ipsec-policy-manual-policy002-10] proposal tran2
[SwitchC-ipsec-policy-manual-policy002-10] sa spi outbound esp 54321
[SwitchC-ipsec-policy-manual-policy002-10] sa spi inbound esp 54321
[SwitchC-ipsec-policy-manual-policy002-10] sa string-key outbound esp gfedcba
[SwitchC-ipsec-policy-manual-policy002-10] sa string-key inbound esp gfedcba
```

```
[SwitchC-ipsec-policy-manual-policy002-10] quit
```

4. Apply the IPsec policies in areas:

Configure Switch A.

```
[SwitchA] ospfv3 1
[SwitchA-ospfv3-1] area 1
[SwitchA-ospfv3-1-area-0.0.0.1] enable ipsec-policy policy001
[SwitchA-ospfv3-1-area-0.0.0.1] quit
[SwitchA-ospfv3-1] quit
```

Configure Switch B.

```
[SwitchB] ospfv3 1
[SwitchB-ospfv3-1] area 0
[SwitchB-ospfv3-1-area-0.0.0.0] enable ipsec-policy policy002
[SwitchB-ospfv3-1-area-0.0.0.0] quit
[SwitchB-ospfv3-1] area 1
[SwitchB-ospfv3-1-area-0.0.0.1] enable ipsec-policy policy001
[SwitchB-ospfv3-1-area-0.0.0.1] quit
[SwitchB-ospfv3-1] quit
```

Configure Switch C.

```
[SwitchC] ospfv3 1
[SwitchC-ospfv3-1] area 0
[SwitchC-ospfv3-1-area-0.0.0.0] enable ipsec-policy policy002
[SwitchC-ospfv3-1-area-0.0.0.0] quit
[SwitchC-ospfv3-1] quit
```

5. Verify the configuration:

OSPFv3 traffic between Switches A, B, and C is protected by IPsec.

Troubleshooting OSPFv3 configuration

No OSPFv3 neighbor relationship established

Symptom

No OSPF neighbor relationship can be established.

Analysis

If the physical link and lower protocol function properly, check OSPF parameters configured on interfaces. The two neighboring interfaces must have the same parameters, such as the area ID, network segment and mask, and network type. If the network type is broadcast, at least one interface must have a DR priority higher than 0.

Process steps

1. Display neighbor information using the **display ospfv3 peer** command.
2. Display OSPFv3 interface information using the **display ospfv3 interface** command.
3. Ping the neighbor router's IP address to check connectivity.
4. Check OSPF timers. The dead interval on an interface must be at least four times the hello interval.
5. On a broadcast network, at least one interface must have a DR priority higher than 0.

Incorrect routing information

Symptom

OSPFv3 cannot find routes to other areas.

Analysis

The backbone area must maintain connectivity to all other areas. If a router connects to more than one area, at least one area must be connected to the backbone. The backbone cannot be configured as a stub area.

In a stub area, routers cannot receive external routes, and interfaces connected to the stub area must be associated with the stub area.

Solution

1. Use the **display ospfv3 peer** command to display OSPFv3 neighbors.
2. Use the **display ospfv3 interface** command to display OSPFv3 interface information.
3. Use the **display ospfv3 lsdb** command to display LSDB information to check integrity.
4. Display information about area configuration using the **display current-configuration configuration** command. If more than two areas are configured, at least one area is connected to the backbone.
5. In a stub area, all routers are configured with the **stub** command.
6. If a virtual link is configured, use the **display ospf vlink** command to check the neighbor state.

Configuring IPv6 IS-IS

This chapter describes how to configure IPv6 IS-IS, which supports all IPv4 IS-IS features except that it advertises IPv6 routing information instead. For information about IS-IS, see "[Configuring IS-IS](#)."

The term "router" in this chapter refers to both routers and Layer 3 switches.

Hardware compatibility

The HP 5500 SI Switch Series does not support IPv6 IS-IS.

Introduction to IPv6 IS-IS

The IS-IS routing protocol (Intermediate System-to-Intermediate System intra-domain routing information exchange protocol) supports multiple network protocols, including IPv6. The international engineer task force (IETF) defines two type-length-values (TLVs) and a new network layer protocol identifier (NLPID) to enable IS-IS to support IPv6.

TLV is a variable-length field in the link state PDU or link state packet (LSP). The TLVs include the following:

- **IPv6 Reachability**—Defines the prefix, metric of routing information to indicate network reachability, and has a type value of 236 (0xEC).
- **IPv6 Interface Address**—Same as the "IP Interface Address" TLV in IPv4 ISIS, except the 32-bit IPv4 address is translated to the 128-bit IPv6 address.

The NLPID is an eight-bit field that identifies which network layer protocol is supported. For IPv6, the NLPID is 142 (0x8E), which must be carried in hello packets sent by a router that supports IPv6 IS-IS.

Configuring IPv6 IS-IS basic functions

You can implement IPv6 inter-networking through configuring IPv6 IS-IS in IPv6 network environment.

Before the configuration, complete the following tasks:

- Enable IPv6 globally.
- Configure IP addresses for interfaces, and ensure all neighboring nodes are reachable.
- Enable IS-IS.

To configure the basic functions of IPv6 IS-IS:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable an IS-IS process and enter IS-IS view.	isis [process-id]	Not enabled by default.
3. Configure the network entity title for the IS-IS process.	network-entity net	Not configured by default.
4. Enable IPv6 for the IS-IS process.	ipv6 enable	Disabled by default.

Step	Command	Remarks
5. Return to system view.	quit	N/A
6. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
7. Enable IPv6 for an IS-IS process on the interface.	isis ipv6 enable [<i>process-id</i>]	Disabled by default.

Configuring IPv6 IS-IS routing information control

Complete the IPv6 IS-IS basic function configuration before configuring this task.

For information about ACL, see *ACL and QoS Configuration Guide*.

For information about routing policy and IPv6 prefix list, see "[Configuring routing policies](#)."

To configure IPv6 IS-IS routing information control:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IS-IS view.	isis [<i>process-id</i>]	N/A
3. Define the priority for IPv6 IS-IS routes.	ipv6 preference { route-policy <i>route-policy-name</i> <i>preference</i> } *	Optional. 15 by default.
4. Configure an IPv6 IS-IS summary route.	ipv6 summary <i>ipv6-prefix</i> <i>prefix-length</i> [avoid-feedback generate_null0_route [level-1 level-1-2 level-2] tag <i>tag</i>] *	Optional. Not configured by default.
5. Generate an IPv6 IS-IS default route.	ipv6 default-route-advertise [[level-1 level-1-2 level-2] route-policy <i>route-policy-name</i>] *	Optional. By default, no IPv6 default route is defined.
6. Configure IPv6 IS-IS to filter incoming routes.	ipv6 filter-policy { <i>acl6-number</i> ipv6-prefix <i>ipv6-prefix-name</i> route-policy <i>route-policy-name</i> } import	Optional. By default, no filtering policy is defined.
7. Configure IPv6 IS-IS to redistribute routes from another routing protocol.	ipv6 import-route <i>protocol</i> [<i>process-id</i>] [allow-ibgp] [cost <i>cost</i> [level-1 level-1-2 level-2] route-policy <i>route-policy-name</i> tag <i>tag</i>] *	Optional. Not configured by default.
8. Configure the maximum number of redistributed Level 1/Level 2 IPv6 routes.	ipv6 import-route limit <i>number</i>	Optional. The default varies with devices.
9. Configure the filtering of outgoing redistributed routes.	ipv6 filter-policy { <i>acl6-number</i> ipv6-prefix <i>ipv6-prefix-name</i> route-policy <i>route-policy-name</i> } export [<i>protocol</i> [<i>process-id</i>]]	Optional. Not configured by default.
10. Enable route leaking.	ipv6 import-route isisv6 level-2 into level-1 [filter-policy { <i>acl6-number</i> ipv6-prefix <i>ipv6-prefix-name</i> route-policy <i>route-policy-name</i> } tag <i>tag</i>] *	Optional. Not enabled by default.

Step	Command	Remarks
11. Specify the maximum number of equal-cost load balanced routes.	ipv6 maximum load-balancing <i>number</i>	Optional. 8 by default.

NOTE:

The **ipv6 filter-policy export** command is usually combined with the **ipv6 import-route** command. If no *protocol* is specified for the **ipv6 filter-policy export** command, routes redistributed from all routing protocols are filtered before advertisement. If a protocol is specified, only routes redistributed from the routing protocol are filtered for advertisement.

Configuring BFD for IPv6 IS-IS

Bidirectional forwarding detection (BFD) provides a mechanism to quickly detect the connectivity of links between IPv6 IS-IS neighbors, thus to improve the convergence speed of IPv6 IS-IS. For more information about BFD, see *High Availability Configuration Guide*.

To configure BFD for IPv6 IS-IS:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable an IS-IS process and enter IS-IS view.	isis [<i>process-id</i>]	N/A
3. Configure the network entity title for the IS-IS process.	network-entity <i>net</i>	Not configured by default.
4. Enable IPv6 for the IS-IS process.	ipv6 enable	Disabled by default.
5. Return to system view.	quit	N/A
6. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
7. Enable IPv6 for an IS-IS process on the interface.	isis ipv6 enable [<i>process-id</i>]	Disabled by default.
8. Enable BFD on the interface.	isis ipv6 bfd enable	Not enabled by default.

Displaying and maintaining IPv6 IS-IS

Task	Command	Remarks
Display brief IPv6 IS-IS information.	display isis brief [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the status of the debug switches.	display isis debug-switches { <i>process-id</i> vpn-instance <i>vpn-instance-name</i> } [[{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display IS-IS enabled interface information.	display isis interface [statistics [<i>interface-type interface-number</i>] [verbose] [<i>process-id</i> vpn-instance vpn-instance-name] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display LSDB information.	display isis lsdb [[I1 I2 level-1 level-2] [[lsp-id lsp-id lsp-name lspname local] verbose] * [<i>process-id</i> vpn-instance vpn-instance-name] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IS-IS mesh group information.	display isis mesh-group [<i>process-id</i> vpn-instance vpn-instance-name] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the mapping table between the host name and system ID.	display isis name-table [<i>process-id</i> vpn-instance vpn-instance-name] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IS-IS neighbor information.	display isis peer [statistics verbose] [<i>process-id</i> vpn-instance vpn-instance-name] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 IS-IS routing information.	display isis route ipv6 [[level-1 level-2] verbose] * [<i>process-id</i> vpn-instance vpn-instance-name] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display SPF log information.	display isis spf-log [<i>process-id</i> vpn-instance vpn-instance-name] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the statistics of the IS-IS process.	display isis statistics [level-1 level-1-2 level-2] [<i>process-id</i> vpn-instance vpn-instance-name] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear all IS-IS data structure information.	reset isis all [<i>process-id</i> vpn-instance vpn-instance-name]	Available in user view
Clear the IS-IS data information of a neighbor.	reset isis peer system-id [<i>process-id</i> vpn-instance vpn-instance-name]	Available in user view

IPv6 IS-IS configuration examples

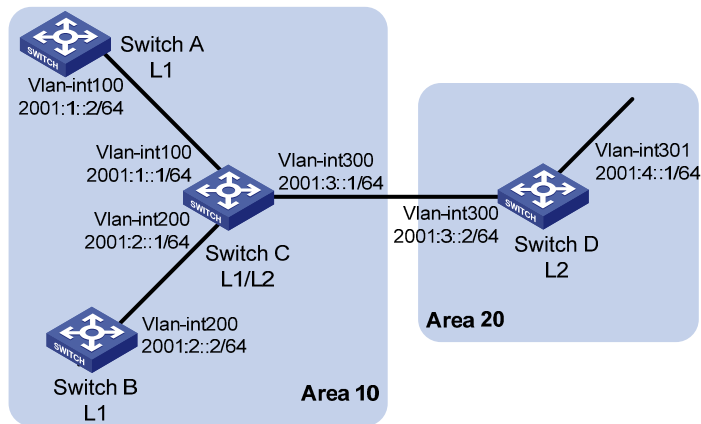
IPv6 IS-IS basic configuration example

Network requirements

As shown in [Figure 113](#), Switch A, Switch B, Switch C, and Switch D reside in the same autonomous system, and all are enabled with IPv6.

Switch A and Switch B are Level-1 switches, Switch D is a Level-2 switch, and Switch C is a Level-1-2 switch. Switch A, Switch B, and Switch C are in area 10, and Switch D is in area 20.

Figure 113 Network diagram



Configuration procedure

1. Configure IPv6 addresses for interfaces. (Details not shown.)
2. Configure IPv6 IS-IS:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] isis 1
[SwitchA-isis-1] is-level level-1
[SwitchA-isis-1] network-entity 10.0000.0000.0001.00
[SwitchA-isis-1] ipv6 enable
[SwitchA-isis-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] isis ipv6 enable 1
[SwitchA-Vlan-interface100] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] isis 1
[SwitchB-isis-1] is-level level-1
[SwitchB-isis-1] network-entity 10.0000.0000.0002.00
[SwitchB-isis-1] ipv6 enable
[SwitchB-isis-1] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] isis ipv6 enable 1
[SwitchB-Vlan-interface200] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] isis 1
[SwitchC-isis-1] network-entity 10.0000.0000.0003.00
[SwitchC-isis-1] ipv6 enable
[SwitchC-isis-1] quit
[SwitchC] interface vlan-interface 100
```

```
[SwitchC-Vlan-interface100] isis ipv6 enable 1
[SwitchC-Vlan-interface100] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] isis ipv6 enable 1
[SwitchC-Vlan-interface200] quit
[SwitchC] interface vlan-interface 300
[SwitchC-Vlan-interface300] isis ipv6 enable 1
[SwitchC-Vlan-interface300] quit
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] ipv6
[SwitchD] isis 1
[SwitchD-isis-1] is-level level-2
[SwitchD-isis-1] network-entity 20.0000.0000.0004.00
[SwitchD-isis-1] ipv6 enable
[SwitchD-isis-1] quit
[SwitchD] interface vlan-interface 300
[SwitchD-Vlan-interface300] isis ipv6 enable 1
[SwitchD-Vlan-interface300] quit
[SwitchD] interface vlan-interface 301
[SwitchD-Vlan-interface301] isis ipv6 enable 1
[SwitchD-Vlan-interface301] quit
```

3. Verify the configuration:

Display the IPv6 IS-IS routing table of Switch A.

```
[SwitchA] display isis route ipv6
```

Route information for ISIS(1)

ISIS(1) IPv6 Level-1 Forwarding Table

Destination: ::	PrefixLen: 0
Flag : R/-/-	Cost : 10
Next Hop : FE80::200:FF:FE0F:4	Interface: Vlan100
Destination: 2001:1::	PrefixLen: 64
Flag : D/L/-	Cost : 10
Next Hop : Direct	Interface: Vlan100
Destination: 2001:2::	PrefixLen: 64
Flag : R/-/-	Cost : 20
Next Hop : FE80::200:FF:FE0F:4	Interface: Vlan100
Destination: 2001:3::	PrefixLen: 64
Flag : R/-/-	Cost : 20
Next Hop : FE80::200:FF:FE0F:4	Interface: Vlan100

Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set

Display the IPv6 IS-IS routing table of Switch B.

[SwitchB] display isis route ipv6

Route information for ISIS(1)

ISIS(1) IPv6 Level-1 Forwarding Table

```
-----  
Destination: ::                               PrefixLen: 0  
Flag       : R/-/-                           Cost      : 10  
Next Hop   : FE80::200:FF:FE0F:4             Interface: Vlan200  
  
Destination: 2001:1::                         PrefixLen: 64  
Flag       : D/L/-                           Cost      : 10  
Next Hop   : FE80::200:FF:FE0F:4             Interface: Vlan200  
  
Destination: 2001:2::                         PrefixLen: 64  
Flag       : R/-/-                           Cost      : 20  
Next Hop   : Direct                          Interface: Vlan200  
  
Destination: 2001:3::                         PrefixLen: 64  
Flag       : R/-/-                           Cost      : 20  
Next Hop   : FE80::200:FF:FE0F:4             Interface: Vlan200
```

Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set

Display the IPv6 IS-IS routing table of Switch C.

[SwitchC] display isis route ipv6

Route information for ISIS(1)

ISIS(1) IPv6 Level-1 Forwarding Table

```
-----  
Destination: 2001:1::                         PrefixLen: 64  
Flag       : D/L/-                           Cost      : 10  
Next Hop   : Direct                          Interface: Vlan100  
  
Destination: 2001:2::                         PrefixLen: 64  
Flag       : D/L/-                           Cost      : 10  
Next Hop   : Direct                          Interface: Vlan200  
  
Destination: 2001:3::                         PrefixLen: 64  
Flag       : D/L/-                           Cost      : 10  
Next Hop   : Direct                          Interface: Vlan300
```

Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set

ISIS(1) IPv6 Level-2 Forwarding Table

```

-----
Destination: 2001:1::                               PrefixLen: 64
Flag       : D/L/-                                  Cost       : 10
Next Hop   : Direct                                 Interface:  Vlan100

Destination: 2001:2::                               PrefixLen: 64
Flag       : D/L/-                                  Cost       : 10
Next Hop   : Direct                                 Interface:  Vlan200

Destination: 2001:3::                               PrefixLen: 64
Flag       : D/L/-                                  Cost       : 10
Next Hop   : Direct                                 Interface:  Vlan300

Destination: 2001:4::1                             PrefixLen: 128
Flag       : R/-/-                                  Cost       : 10
Next Hop   : FE80::20F:E2FF:FE3E:FA3D              Interface:  Vlan300

```

Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set

Display the IPv6 IS-IS routing table of Switch D.

[SwitchD] display isis route ipv6

Route information for ISIS(1)

ISIS(1) IPv6 Level-2 Forwarding Table

```

-----
Destination: 2001:1::                               PrefixLen: 64
Flag       : R/-/-                                  Cost       : 20
Next Hop   : FE80::200:FF:FE0F:4                  Interface:  Vlan300

Destination: 2001:2::                               PrefixLen: 64
Flag       : R/-/-                                  Cost       : 20
Next Hop   : FE80::200:FF:FE0F:4                  Interface:  Vlan300

Destination: 2001:3::                               PrefixLen: 64
Flag       : D/L/-                                  Cost       : 10
Next Hop   : Direct                                 Interface:  Vlan300

Destination: 2001:4::1                             PrefixLen: 128
Flag       : D/L/-                                  Cost       : 0
Next Hop   : Direct                                 Interface:  Loop1

```

Flags: D-Direct, R-Added to RM, L-Advertised in LSPs, U-Up/Down Bit Set

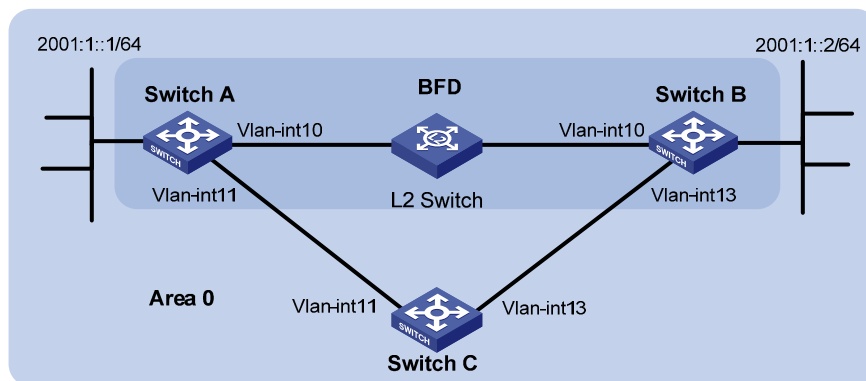
BFD for IPv6 IS-IS configuration example

Network requirements

As shown in [Figure 114](#), configure IPv6 IS-IS on Switch A, Switch B, and Switch C and configure BFD over the link Switch A \leftrightarrow L2 Switch \leftrightarrow Switch B.

After the link between Switch B and the Layer-2 switch fails, BFD can quickly detect the failure and notify IPv6 IS-IS of the failure. Then Switch A and Switch B communicate through Switch C.

Figure 114 Network diagram



Device	Interface	IPv6 address	Device	Interface	IPv6 address
Switch A	Vlan-int10	2001::1/64	Switch B	Vlan-int10	2001::2/64
	Vlan-int11	2001:2::1/64		Vlan-int13	2001:3::2/64
Switch C	Vlan-int11	2001:2::2/64			
	Vlan-int13	2001:3::1/64			

Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)

2. Configure IPv6 IS-IS:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] isis 1
[SwitchA-isis-1] is-level level-1
[SwitchA-isis-1] network-entity 10.0000.0000.0001.00
[SwitchA-isis-1] ipv6 enable
[SwitchA-isis-1] quit
[SwitchA] interface vlan-interface 10
[SwitchA-Vlan-interface10] isis ipv6 enable 1
[SwitchA-Vlan-interface10] quit
[SwitchA] interface vlan-interface 11
[SwitchA-Vlan-interface11] isis ipv6 enable 1
[SwitchA-Vlan-interface11] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] isis 1
```

```

[SwitchB-isis-1] is-level level-1
[SwitchB-isis-1] network-entity 10.0000.0000.0002.00
[SwitchB-isis-1] ipv6 enable
[SwitchB-isis-1] quit
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] isis ipv6 enable 1
[SwitchB-Vlan-interface10] quit
[SwitchB] interface vlan-interface 13
[SwitchB-Vlan-interface13] isis ipv6 enable 1
[SwitchB-Vlan-interface13] quit

```

Configure Switch C.

```

<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] isis 1
[SwitchC-isis-1] network-entity 10.0000.0000.0003.00
[SwitchC-isis-1] ipv6 enable
[SwitchC-isis-1] quit
[SwitchC] interface vlan-interface 11
[SwitchC-Vlan-interface11] isis ipv6 enable 1
[SwitchC-Vlan-interface11] quit
[SwitchC] interface vlan-interface 13
[SwitchC-Vlan-interface13] isis ipv6 enable 1
[SwitchC-Vlan-interface13] quit

```

3. Configure BFD functions:

Enable BFD on Switch A and configure BFD parameters.

```

[SwitchA] bfd session init-mode active
[SwitchA] interface vlan-interface 10
[SwitchA-Vlan-interface10] isis ipv6 bfd enable
[SwitchA-Vlan-interface10] bfd min-transmit-interval 500
[SwitchA-Vlan-interface10] bfd min-receive-interval 500
[SwitchA-Vlan-interface10] bfd detect-multiplier 7
[SwitchA-Vlan-interface10] return

```

Enable BFD on Switch B and configure BFD parameters.

```

[SwitchB] bfd session init-mode active
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] isis ipv6 bfd enable
[SwitchB-Vlan-interface10] bfd min-transmit-interval 500
[SwitchB-Vlan-interface10] bfd min-receive-interval 500
[SwitchB-Vlan-interface10] bfd detect-multiplier 6

```

4. Verify configuration:

The following operations are made on Switch A. Operations for Switch B are similar and are not shown here.

Display the BFD information of Switch A.

```

<SwitchA> display bfd session
Total Session Num: 1          Init Mode: Active

IPv6 Session Working Under Ctrl Mode:

```

```

Local Discr: 1441 Remote Discr: 1450
Source IP: FE80::20F:FF:FE00:1202 (link-local address of VLAN-interface 10 on
Switch A)
Destination IP: FE80::20F:FF:FE00:1200 (link-local address of VLAN-interface 10 on
Switch B)
Session State: Up Interface: Vlan10
Hold Time: /

```

Display route 2001:4::0/64 on Switch A, and you can see that Switch A and Switch B communicate through the Layer-2 switch.

```
<SwitchA> display ipv6 routing-table 2001:4::0 64 verbose
```

```
Routing Table :
```

```
Summary Count : 2
```

```

Destination : 2001:4::0 PrefixLength : 64
NextHop : 2001::2 Preference : 15
IpPrecedence : QoS LcId :
RelayNextHop : :: Tag : 0H
Neighbor : :: ProcessID : 0
Interface : Vlan-interface10 Protocol : ISISv6
State : Active Adv Cost : 20
Tunnel ID : 0x0 Label : NULL
Age : 4538sec

```

```

Destination : 2001:4::0 PrefixLength : 64
NextHop : 2001:2::2 Preference : 15
IpPrecedence : QoS LcId :
RelayNextHop : :: Tag : 0H
Neighbor : :: ProcessID : 0
Interface : Vlan-interface11 Protocol : ISISv6
State : Invalid Adv Cost : 30
Tunnel ID : 0x0 Label : NULL
Age : 4515sec

```

Enable BFD debugging on Switch A.

```

<SwitchA> debugging bfd scm
<SwitchA> debugging bfd event
<SwitchA> debugging isis event bfd
<SwitchA> terminal debugging

```

When the link between Switch B and the Layer-2 switch fails, BFD can quickly detect the failure.

```

#Aug 8 14:54:05:362 2009 SwitchA IFNET/4/INTERFACE UPDOWN:
Trap 1.3.6.1.6.3.1.1.5.3<linkDown>: Interface 983041 is Down, ifAdminStatus is
1, ifOperStatus is 2
#Aug 8 14:54:05:363 2009 SwitchA ISIS/4/ADJ_CHANGE:TrapID(1.3.6.1.2.1.138.0.17<
isisAdjacencyChange>), ISIS Level-2 Adjacency IN Circuit-983041 State Change.
#Aug 8 14:54:05:364 2009 SwitchA ISIS/4/ADJ_CHANGE:TrapID(1.3.6.1.2.1.138.0.17<
isisAdjacencyChange>), ISIS Level-1 Adjacency IN Circuit-983041 State Change.
%Aug 8 14:54:05:365 2009 SwitchA IFNET/4/LINK UPDOWN: vlan10: link status is DOWN

```

%Aug 8 14:54:05:366 2009 SwitchA IFNET/4/UPDOWN: Line protocol on the interface Vlan10 is DOWN

%Aug 8 14:54:05:367 2009 SwitchA ISIS/4/ADJLOG:ISIS-1-ADJCHANGE: Adjacency To 0000.0000.0002 (vlan10) DOWN, Level-2 Circuit Down.

%Aug 8 14:54:05:367 2009 SwitchA ISIS/4/ADJLOG:ISIS-1-ADJCHANGE: Adjacency To 0000.0000.0002 (vlan10) DOWN, Level-2 Adjacency clear.

%Aug 8 14:54:05:368 2009 SwitchA ISIS/4/ADJLOG:ISIS-1-ADJCHANGE: Adjacency To 0000.0000.0002 (vlan10) DOWN, Level-1 Circuit Down.

%Aug 8 14:54:05:369 2009 SwitchA ISIS/4/ADJLOG:ISIS-1-ADJCHANGE: Adjacency To 0000.0000.0002 (vlan10) DOWN, Level-1 Adjacency clear.

*Aug 8 14:54:05:369 2009 SwitchA ISIS/6/ISIS: ISIS-1-BFD: Recieve BFD session down . Type 0. DstIPAddr: FE80::20F:FF:FE00:1200 , SrcIPAddr: FE80::20F:FF:FE00:1202

*Aug 8 14:54:05:370 2009 SwitchA ISIS/6/ISIS: ISIS-1-BFD: Success to send msg. Msg type 1 delete session. IfPhyIndex: 5 , DstIPAddr: FE80::20F:FF:FE00:1200 , SrcIPAddr: FE80::20F:FF:FE00:1202. NeighborType:Level-1.

Display the BFD information of Switch A. You can see that Switch A has removed its neighbor relationship with Switch B and therefore no information is output.

<SwitchA> display bfd session

Display route 2001:4::0/64 on Switch A, and you can see that Switch A and Switch B communicate through Switch C.

<SwitchA> display ipv6 routing-table 2001:4::0 64 verbose

Routing Table :

Summary Count : 1

Destination	: 2001:4::0	PrefixLength	: 64
NextHop	: 2001:2::2	Preference	: 15
IpPrecedence	:	QoS LcId	:
RelayNextHop	: ::	Tag	: 0H
Neighbor	: ::	ProcessID	: 0
Interface	: Vlan-interface11	Protocol	: ISISv6
State	: Invalid Adv	Cost	: 30
Tunnel ID	: 0x0	Label	: NULL
Age	: 4610sec		

Configuring IPv6 BGP

This chapter describes only configuration for IPv6 BGP. For BGP related information, see "[Configuring BGP](#)." The term "router" in this chapter refers to both routers and Layer 3 switches.

Hardware compatibility

The HP 5500 SI Switch Series does not support IPv6 BGP.

IPv6 BGP overview

BGP-4 was designed to carry only IPv4 routing information, so other network layer protocols such as IPv6 are not supported.

To support multiple network layer protocols, IETF extended BGP-4 by introducing Multiprotocol BGP (MP-BGP), which is defined in RFC 2858 (multiprotocol extensions for BGP-4).

For brevity purposes, MP-BGP for IPv6 is called "IPv6 BGP."

IPv6 BGP puts IPv6 network layer information into the attributes of Network Layer Reachability Information (NLRI) and NEXT_HOP.

The NLRI attribute of IPv6 BGP involves the following:

- **MP_REACH_NLRI**—Multiprotocol Reachable NLRI, for advertising reachable route and next hop information.
- **MP_UNREACH_NLRI**—Multiprotocol Unreachable NLRI, for withdrawal of unreachable routes.

The NEXT_HOP attribute of IPv6 BGP is identified by an IPv6 unicast address or IPv6 local link address.

IPv6 BGP has the same messaging and routing mechanisms as BGP.

IPv6 BGP configuration task list

Task	Remarks
Configuring IPv6 BGP basic Specifying an IPv6 BGP peer	Required

Task	Remarks	
	Injecting a local IPv6 route	Optional
	Configuring a preferred value for routes from a peer or peer group	Optional
	Specifying the source interface for establishing TCP connections	Optional
	Allowing the establishment of an indirect EBGp connection	Optional
	Configuring a description for an IPv6 peer or peer group	Optional
	Disabling session establishment to an IPv6 peer or peer group	Optional
	Logging IPv6 peer or peer group state changes	Optional
Controlling route distribution and reception	Configuring IPv6 BGP route redistribution	Optional
	Configuring IPv6 BGP route summarization	Optional
	Advertising a default route to an IPv6 peer or peer group	Optional
	Configuring outbound route filtering	Optional
	Configuring inbound route filtering	Optional
	Configuring IPv6 BGP and IGP route synchronization	Optional
Configuring IPv6 BGP route attributes	Configuring route dampening	Optional
	Configuring IPv6 BGP preference and default LOCAL_PREF and NEXT_HOP attributes	Optional
	Configuring the MED attribute	Optional
Tuning and optimizing IPv6 BGP networks	Configuring the AS_PATH attribute	Optional
	Configuring IPv6 BGP timers	Optional
	Configuring IPv6 BGP soft reset	Optional
	Enabling the IPv6 BGP ORF capability	Optional
	Enabling 4-byte AS number suppression	Optional
	Setting the DSCP value for IPv6 BGP packets	Optional
	Configuring the maximum number of load-balanced routes	Optional
	Enabling MD5 authentication for TCP connections	Optional
Applying an IPsec policy to an IPv6 BGP peer or peer group	Optional	
Configuring a large-scale IPv6 BGP network	Configuring IPv6 BGP peer group	Optional
	Configuring IPv6 BGP community	Optional
	Configuring an IPv6 BGP route reflector	Optional

Task	Remarks
Configuring BFD for IPv6 BGP	Optional

Configuring IPv6 BGP basic functions

Configuration prerequisites

Before you configure IPv6 BGP basic functions, complete the following tasks:

- Specify IP addresses for interfaces.
- Enable IPv6.

Configuration guidelines

Create a peer group before configuring basic functions for it. For more information, see "[Configuring IPv6 BGP peer group](#)."

Specifying an IPv6 BGP peer

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Specify a router ID.	router-id <i>router-id</i>	Optional. Required if no IP addresses are configured for any interfaces.
4. Enter IPv6 address family view.	ipv6-family	N/A
5. Specify an IPv6 peer.	peer <i>ipv6-address</i> as-number <i>as-number</i>	N/A

Injecting a local IPv6 route

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Inject a local route into the IPv6 BGP routing table.	network <i>ipv6-address prefix-length</i> [short-cut route-policy <i>route-policy-name</i>]	Not added by default.

Configuring a preferred value for routes from a peer or peer group

If you both reference a routing policy and use the command **peer { ipv6-group-name | ipv6-address } preferred-value value** to set a preferred value for routes from a peer, the routing policy sets the specific preferred value for routes matching it. If the preferred value in the routing policy is zero, the routes use the value set with the **peer { ipv6-group-name | ipv6-address } preferred-value value** command. For how to use a routing policy to set a preferred value, see the command **peer { group-name | ipv4-address | ipv6-address } route-policy route-policy-name { import | export }** in this chapter, and the command **apply preferred-value preferred-value** in *Layer 3—IP Routing Command Reference*.

To configure a preferred value for routes from a peer or peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Configure a preferred value for routes received from an IPv6 peer or peer group.	peer { ipv6-group-name ipv6-address } preferred-value value	By default, the preferred value is 0.

Specifying the source interface for establishing TCP connections

IPv6 BGP uses TCP as the transport layer protocol. By default, IPv6 BGP uses the output interface of the optimal route to a peer or peer group as the source interface for establishing TCP connections to the peer or peer group.

If an IPv6 BGP router has multiple links to a peer, and the source interface fails, IPv6 BGP has to reestablish TCP connections, causing network oscillation. To enhance stability of IPv6 BGP connections, HP recommends using a loopback interface as the source interface.

To establish a BGP connection, specify on the local router the source interface for establishing the TCP connection to the peer on the peering BGP router. Otherwise, the local BGP router may fail to establish TCP connection to the peer when using the outbound interface of the best route as the source interface.

To specify the source interface for establishing TCP connections to a BGP peer or peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A

Step	Command	Remarks
4. Specify the source interface for establishing TCP connections to an IPv6 BGP peer or peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } connect-interface <i>interface-type interface-number</i>	By default, IPv6 BGP uses the outbound interface of the best route to the IPv6 BGP peer or peer group as the source interface for establishing a TCP connection.

Allowing the establishment of an indirect EBGP connection

Direct links must be available between EBGP peers. If not, you can use the **peer ebgp-max-hop** command to establish a multi-hop TCP connection in between. However, do not use this command for direct EBGP connections with loopback interfaces.

To allow the establishment of a non-direct EBGP connection:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Allow the establishment of EBGP connection to an indirectly connected peer or peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } ebgp-max-hop [<i>hop-count</i>]	Not configured by default.

Configuring a description for an IPv6 peer or peer group

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Configure a description for an IPv6 peer or peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } description <i>description-text</i>	Optional. Not configured by default. The peer group to be configured with a description must have been created.

Disabling session establishment to an IPv6 peer or peer group

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A

Step	Command	Remarks
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Disable session establishment to an IPv6 peer or peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } ignore	Optional. Not disabled by default.

Logging IPv6 peer or peer group state changes

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enable logging of peer changes globally.	log-peer-change	Optional. Enabled by default. For information about the log-peer-change command, see <i>Layer 3—IP Routing Command Reference</i> .
4. Enter IPv6 address family view.	ipv6-family	N/A
5. Enable the state change logging for an IPv6 peer or peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } log-change	Optional. Enabled by default.

Controlling route distribution and reception

This task includes routing information filtering, routing policy application, and route dampening.

Configuration prerequisites

Before you configure route distribution and reception control, complete the following tasks:

- Enable IPv6.
- Configure IPv6 BGP basic functions.

Configuring IPv6 BGP route redistribution

To configure IPv6 BGP route redistribution:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A

Step	Command	Remarks
4. Enable default route redistribution into the IPv6 BGP routing table.	default-route imported	Optional. Not enabled by default. If the default-route imported command is not configured, using the import-route command cannot redistribute any IGP default route.
5. Enable route redistribution from another routing protocol.	import-route protocol [<i>process-id</i> [med <i>med-value</i> route-policy <i>route-policy-name</i>] *]	Not enabled by default.

Configuring IPv6 BGP route summarization

To reduce the routing table size on medium and large BGP networks, configure route summarization on BGP routers. BGP supports only manual summarization of IPv6 routes.

To configure IPv6 BGP route summarization:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Configure manual route summarization.	aggregate <i>ipv6-address prefix-length</i> [as-set attribute-policy <i>route-policy-name</i> detail-suppressed origin-policy <i>route-policy-name</i> suppress-policy <i>route-policy-name</i>] *	Not configured by default.

Advertising a default route to an IPv6 peer or peer group

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A

Step	Command	Remarks	
4.	Advertise a default route to an IPv6 peer or peer group.	<pre>peer { ipv6-group-name ipv6-address } default-route-advertise [route-policy route-policy-name]</pre>	<p>Not advertised by default.</p> <p>With the peer default-route-advertise command executed, the local router advertises a default route with itself as the next hop to the specified IPv6 peer or peer group, regardless of whether the default route is available in the routing table.</p>

Configuring outbound route filtering

IPv6 BGP advertises routes passing the specified policy to peers. Using the *protocol* argument can filter only the routes redistributed from the specified protocol. If no *protocol* is specified, IPv6 BGP filters all routes to be advertised, including redistributed routes and routes imported with the **network** command.

To configure outbound route filtering:

Step	Command	Remarks	
1.	Enter system view.	system-view	N/A
2.	Enter BGP view.	bgp as-number	N/A
3.	Enter IPv6 address family view.	ipv6-family	N/A
4.	Configure the filtering of outgoing routes.	<pre>filter-policy { acl6-number ipv6-prefix ipv6-prefix-name } export [protocol process-id]</pre>	Not configured by default.
5.	Apply a routing policy to routes advertised to an IPv6 peer or peer group.	<pre>peer { ipv6-group-name ipv6-address } route-policy route-policy-name export</pre>	Not applied by default.
6.	Specify an IPv6 ACL to filter routes advertised to an IPv6 peer or peer group.	<pre>peer { ipv6-group-name ipv6-address } filter-policy acl6-number export</pre>	Not specified by default.
7.	Specify an AS path ACL to filter routes advertised to an IPv6 peer or peer group.	<pre>peer { ipv6-group-name ipv6-address } as-path-acl as-path-acl-number export</pre>	Not specified by default.
8.	Specify an IPv6 prefix list to filter routes advertised to an IPv6 peer or peer group.	<pre>peer { ipv6-group-name ipv6-address } ipv6-prefix ipv6-prefix-name export</pre>	Not specified by default.

Configuring inbound route filtering

Only routes passing the configured filtering can be added into the local IPv6 BGP routing table.

Members of a peer group can have different inbound route filtering policies.

To configure inbound route filtering:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Configure inbound route filtering.	filter-policy { <i>acl6-number</i> ipv6-prefix <i>ipv6-prefix-name</i> } import	Not configured by default.
5. Apply a routing policy to routes from an IPv6 peer or peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } route-policy <i>route-policy-name</i> import	Not applied by default.
6. Specify an ACL to filter routes imported from an IPv6 peer or peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } filter-policy <i>acl6-number</i> import	Not specified by default.
7. Specify an AS path ACL to filter routing information imported from an IPv6 peer or peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } as-path-acl <i>as-path-acl-number</i> import	Not specified by default.
8. Specify an IPv6 prefix list to filter routing information imported from an IPv6 peer or peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } ipv6-prefix <i>ipv6-prefix-name</i> import	Not specified by default.
9. Specify the upper limit of prefixes allowed to receive from an IPv6 peer or peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } route-limit <i>limit</i> [<i>percentage</i>]	Optional. Unlimited by default.

Configuring IPv6 BGP and IGP route synchronization

By default, upon receiving an IBGP route, an IPv6 BGP router checks the route's next hop. If the next hop is reachable, the IPv6 BGP router advertises the route to EBGP peers. If the synchronization feature is configured, in addition to the reachability check of the next hop, the IPv6 BGP router must find an active IGP route with the same destination network segment before it can advertise the IBGP route (use the **display ipv6 routing-table protocol** command to check the IGP route state).

To configure IPv6 BGP and IGP route synchronization:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Enable route synchronization between IPv6 BGP and IGP.	synchronization	Not enabled by default.

Configuring route dampening

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Configure IPv6 BGP route dampening parameters.	dampening [<i>half-life-reachable</i> <i>half-life-unreachable</i> <i>reuse suppress ceiling</i> route-policy <i>route-policy-name</i>]*	Optional. Not configured by default.

Configuring IPv6 BGP route attributes

Use the following IPv6 BGP route attributes to modify BGP routing policy:

- IPv6 BGP protocol preference
- Default LOCAL_PREF attribute
- MED attribute
- NEXT_HOP attribute
- AS_PATH attribute

Configuration prerequisites

Before you configure IPv6 BGP route attributes, complete the following tasks:

- Enable IPv6 function.
- Configure IPv6 BGP basic functions.

Configuring IPv6 BGP preference and default LOCAL_PREF and NEXT_HOP attributes

Follow these guidelines when you configure IPv6 BGP preference and default LOCAL_PREF and NEXT_HOP attributes:

- To ensure an IBGP peer can find the correct next hop, configure routes advertised to the IPv6 IBGP peer or peer group to use the local router as the next hop. If BGP load balancing is configured, the local router specifies itself as the next hop of routes sent to an IPv6 IBGP peer or peer group regardless of whether the **peer next-hop-local** command is configured.
- In a *third party next hop* network where the two IPv6 EBGP peers reside in a common broadcast subnet, the router does not change the next hop for routes sent to the IPv6 EBGP peer or peer group by default, unless the **peer next-hop-local** command is configured.

To configure IPv6 BGP preference and default LOCAL_PREF and NEXT_HOP attributes:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Configure preference values for IPv6 BGP external, internal, and local routes.	preference { <i>external-preference</i> <i>internal-preference</i> <i>local-preference</i> route-policy <i>route-policy-name</i> }	Optional. The default preference values of external, internal, and local routes are 255, 255, and 130.
5. Configure the default local preference.	default local-preference <i>value</i>	Optional. The <i>value</i> defaults to 100.
6. Advertise routes to an IPv6 peer or peer group with the local router as the next hop.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } next-hop-local	By default, IPv6 BGP specifies the local router as the next hop for routes sent to an IPv6 EBGP peer or peer group, but does not change the next hop for routes sent to an IPv6 IBGP peer or peer group.

Configuring the MED attribute

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Configure a default MED value.	default med <i>med-value</i>	Optional. Defaults to 0.
5. Enable the comparison of MED for routes from different EBGP peers.	compare-different-as-med	Optional. Not enabled by default.
6. Enable the comparison of MED for routes from each AS.	bestroute compare-med	Optional. Disabled by default.
7. Enable the comparison of MED for routes from confederation peers.	bestroute med-confederation	Optional. Disabled by default.

Configuring the AS_PATH attribute

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A

Step	Command	Remarks
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Allow the local AS number to appear in AS_PATH of routes from a peer or peer group and specify the repeat times.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } allow-as-loop [<i>number</i>]	Optional. Not allowed by default.
5. Specify a fake AS number for an IPv6 peer or peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } fake-as <i>as-number</i>	Optional. Not specified by default.
6. Disable IPv6 BGP from considering the AS_PATH during best route selection.	bestroute as-path-neglect	Optional. Enabled by default.
7. Configure to carry only the public AS number in updates sent to a peer or peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } public-as-only	Optional. By default, IPv6 BGP updates carry a private AS number.
8. Substitute the local AS number for the AS number of an IPv6 peer or peer group identified in the AS_PATH attribute.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } substitute-as	Optional. Not substituted by default.

Tuning and optimizing IPv6 BGP networks

This section describes configurations of IPv6 BGP timers, IPv6 BGP connection soft reset, and the maximum number of load balanced routes.

- IPv6 BGP timers

After establishing an IPv6 BGP connection, two routers send keepalive messages periodically to each other to maintain the connection. If a router receives no keepalive message from the peer after the holdtime elapses, it tears down the connection.

When establishing an IPv6 BGP connection, the two parties compare their holdtimes, taking the shorter one as the common holdtime. If the holdtime is 0, neither keepalive message is sent, nor holdtime is checked.

- IPv6 BGP connection soft reset

After modifying a route selection policy, you must reset IPv6 BGP connections to make the new one take effect. The current IPv6 BGP implementation supports the route-refresh feature that enables dynamic route refresh without needing to disconnect IPv6 BGP links.

After this feature is enabled on all IPv6 BGP routers, a router that wants to apply a new route selection policy advertises a route-refresh message to its peers, which then send their routing information to the router. After receiving the routing information, the router can perform dynamic route update by using the new policy without tearing down connections.

If a peer not supporting route-refresh exists in the network, you must configure the **peer keep-all-routes** command to save all routes from the peer. When the routing policy is changed, the system will update the IPv6 BGP routing table and apply the new policy.

Configuration prerequisites

Before you configure IPv6 BGP timers, complete the following tasks:

- Enable IPv6.
- Configure IPv6 BGP basic functions.

Configuring IPv6 BGP timers

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Configure IPv6 BGP timers.	<ul style="list-style-type: none">• Specify keepalive interval and holdtime: timer keepalive keepalive hold holdtime• Configure keepalive interval and holdtime for an IPv6 peer or peer group: peer { ipv6-group-name .ipv6-address } timer keepalive keepalive hold holdtime	<p>Optional.</p> <p>The keepalive interval defaults to 60 seconds, holdtime defaults to 180 seconds.</p> <p>The holdtime interval must be at least three times the keepalive interval.</p> <p>Timers configured by using the timer command have lower priority than timers configured by using the peer timer command.</p>
5. Configure the interval for sending the same update to an IPv6 peer or peer group.	peer { ipv6-group-name ipv6-address } route-update-interval interval	<p>Optional.</p> <p>The interval for sending the same update to an IBGP peer or an EBGP peer defaults to 15 seconds or 30 seconds.</p>

Configuring IPv6 BGP soft reset

Enabling route refresh

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Enable route refresh.	peer { ipv6-group-name ipv6-address } capability-advertise route-refresh	<p>Optional.</p> <p>Enabled by default.</p>

Performing manual soft-reset

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Save all routes from an IPv6 peer or peer group, not letting them go through the inbound policy.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } keep-all-routes	Optional. Not saved by default. If the peer keep-all-routes command is used, all routes from the peer or peer group are saved regardless of whether the filtering policy is available. These routes will be used to generate IPv6 BGP routes after soft-reset is performed.
5. Return to user view.	return	N/A
6. Soft-reset BGP connections manually.	refresh bgp ipv6 { all <i>ipv6-address</i> group <i>ipv6-group-name</i> external internal } { export import }	N/A

Enabling the IPv6 BGP ORF capability

The BGP Outbound Route Filter (ORF) feature allows a BGP speaker to send its BGP peer a set of ORFs through route-refresh messages. The peer then applies the ORFs, in addition to its local routing policies (if any), to filter updates to the BGP speaker, reducing the number of exchanged update messages and saving network resources.

After you enable the BGP ORF capability, the local BGP router negotiates the ORF capability with the BGP peer through Open messages. The local BGP router determines whether to carry ORF information in messages. If yes, it will further determine whether to carry non-standard ORF information in the packets. After completing the negotiation process and establishing the neighboring relationship, the BGP router and its BGP peer can exchange ORF information through specific route-refresh messages.

For the parameters configured on both sides for ORF capability negotiation, see [Table 9](#).

To enable the BGP ORF capability:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	Required
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Enable BGP route refresh for a peer or peer group.	peer { <i>group-name</i> <i>ipv6-address</i> } capability-advertise route-refresh	Enabled by default

Step	Command	Remarks
5. Enable the non-standard ORF capability for a BGP peer or peer group.	peer { <i>group-name</i> <i>ipv6-address</i> } capability-advertise orf non-standard	Optional. By default, standard BGP ORF capability defined in RFC 5291 and RFC 5292 is supported.
6. Enable the ORF IP prefix negotiation capability for a BGP peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> <i>ipv6-address</i> } capability-advertise orf ip-prefix { both receive send }	Not supported by default.

Table 9 Description of the both, send, and receive parameters and the negotiation result

Local parameter	Peer parameter	Negotiation result
send	<ul style="list-style-type: none"> • receive • both 	The ORF sending capability is enabled locally and the ORF receiving capability is enabled on the peer.
receive	<ul style="list-style-type: none"> • send • both 	The ORF receiving capability is enabled locally and the ORF sending capability is enabled on the peer.
both	both	Both the ORF sending and receiving capabilities are enabled locally and on the peer.

Enabling 4-byte AS number suppression

When a switch that supports 4-byte AS numbers sends an Open message for peer relationship establishment, the Optional parameters field of the message indicates that the AS number occupies four bytes—in the range of 1 to 4294967295. If the peer device does not support 4-byte AS numbers (for examples, it supports only 2-byte AS numbers), the peer relationship cannot be established.

After you enable the 4-byte AS number suppression function, the peer device can then process the Open message even though it does not support 4-byte AS numbers, and the BGP peer relationship can be established.

If the peer device supports 4-byte AS numbers, do not enable the 4-byte AS number suppression function; otherwise, the BGP peer relationship cannot be established.

To enable 4-byte AS number suppression:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Enable 4-byte AS number suppression.	peer { <i>group-name</i> <i>ip-address</i> } capability-advertise suppress-4-byte-as	Disabled by default.

Setting the DSCP value for IPv6 BGP packets

An IPv6 packet header contains an 8-bit Traffic class field. This field identifies the service type of IPv6 packets. As defined in RFC 2474, the first six bits set the Differentiated Services Code Point (DSCP) value and the last two bits are reserved. Network devices use the DSCP value as a reference to determine the packet priority for transmission.

You can set the DSCP value for IPv6 BGP packets.

To set the DSCP value for packets sent to an IPv6 BGP peer or peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Set the DSCP value for the BGP packets sent to the specified IPv6 peer or peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } dscp <i>dscp-value</i>	Optional. By default, the DSCP value in IPv6 BGP packets is 48.

Configuring the maximum number of load-balanced routes

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Configure the maximum number of load balanced routes.	balance <i>number</i>	By default, no load balancing is enabled.

Enabling MD5 authentication for TCP connections

IPv6 BGP employs TCP as the transport protocol. To enhance security, configure IPv6 BGP to perform MD5 authentication when establishing a TCP connection. If the authentication fails, no TCP connection can be established.

The MD5 authentication for establishing TCP connections does not apply to BGP packets.

The MD5 authentication requires that the two parties have the same authentication mode and password to establish a TCP connection; otherwise, no TCP connection can be established due to authentication failure.

To enable MD5 authentication for TCP connections:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Enable MD5 authentication when establishing a TCP connection to the peer or peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } password { cipher simple } <i>password</i>	Not enabled by default.

Applying an IPsec policy to an IPv6 BGP peer or peer group

To protect routing information and defend attacks, IPv6 BGP can authenticate protocol packets by using an IPsec policy.

Outbound IPv6 BGP packets carry the Security Parameter Index (SPI) defined in the IPsec policy. A device uses the SPI carried in a received packet to match against the configured IPsec policy. If they match, the device accepts the packet; otherwise, it discards the packet and will not establish a neighbor relationship with the sending device.

Configuration prerequisites

Before you apply an IPsec policy to a peer or peer group, complete following tasks:

- Create an IPsec proposal.
- Create an IPsec policy.

For more information about IPsec policy configuration, see *Security Configuration Guide*.

Configuration guidelines

An IPsec policy used for IPv6 BGP can be only in manual mode. For more information, see *Security Configuration Guide*.

Configuration procedure

To apply an IPsec policy to a peer or peer group

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Apply an IPsec policy to a peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> } ipsec-policy <i>policy-name</i>	Not configured by default.

Configuring a large-scale IPv6 BGP network

In a large-scale IPv6 BGP network, configuration and maintenance become inconvenient because of too many peers. Configuring peer groups makes management easier and improves route distribution efficiency. Peer group includes IBGP peer group, where peers belong to the same AS, and EBGP peer group, where peers belong to different ASs. If peers in an EBGP group belong to the same external AS, the EBGP peer group is a pure EBGP peer group, and if not, a mixed EBGP peer group.

In a peer group, all members have a common policy. Using the community attribute can make a set of IPv6 BGP routers in multiple ASs have the same policy because community sending between IPv6 BGP peers is not limited by AS.

To assure connectivity between IBGP peers, make them fully meshed, but it becomes impractical when too many IBGP peers exist. Using route reflectors or confederation can solve this issue. In a large-scale AS, both of them can be used.

Confederation configuration of IPv6 BGP is identical to that of BGP4, so it is not mentioned here.

Configuration prerequisites

Before you configure a large-scale IPv6 BGP network, complete the following tasks:

- Make peer nodes accessible to each other at the network layer.
- Enable BGP and configure a router ID.

Configuring IPv6 BGP peer group

Configuring an IBGP peer group

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Create an IBGP peer group.	group ipv6-group-name [internal]	N/A
5. Add a peer into the group.	peer ipv6-address group ipv6-group-name [as-number as-number]	Not added by default.

Creating a pure EBGP peer group

To create a pure EBGP peer group, specify an AS number for the peer group.

If a peer was added into an EBGP peer group, you cannot specify any AS number for the peer group.

To create a pure EBGP peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Create an EBGp peer group.	group <i>ipv6-group-name</i> external	N/A
5. Configure the AS number for the peer group.	peer <i>ipv6-group-name</i> as-number <i>as-number</i>	Not configured by default.
6. Add an IPv6 peer into the peer group.	peer <i>ipv6-address</i> group <i>ipv6-group-name</i>	Not added by default.

Creating a mixed EBGp peer group

When creating a mixed EBGp peer group, you must create a peer and specify its AS number that is different from AS numbers of other peers; however, you cannot specify AS number for the EBGp peer group.

To create a mixed EBGp peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Create an EBGp peer group.	group <i>ipv6-group-name</i> external	N/A
5. Specify the AS number of an IPv6 peer.	peer <i>ipv6-address</i> as-number <i>as-number</i>	Not specified by default.
6. Add the IPv6 peer into the peer group.	peer <i>ipv6-address</i> group <i>ipv6-group-name</i>	Not added by default.

Configuring IPv6 BGP community

When configuring IPv6 BGP community, you must configure a routing policy to define the community attribute, and apply the routing policy to route advertisement. For routing policy configuration, see "[Configuring routing policies.](#)"

Advertising community attribute to an IPv6 peer or peer group

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A

Step	Command	Remarks
4. Advertise community attribute to an IPv6 peer or peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } advertise-community	Not advertised by default.
5. Advertise extended community attribute to an IPv6 peer or peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } advertise-ext-community	Not advertised by default.

Applying a routing policy to routes advertised to a peer or peer group

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Apply a routing policy to routes advertised to an IPv6 peer or peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } route-policy <i>route-policy-name</i> export	Not applied by default.

Configuring an IPv6 BGP route reflector

Follow these guidelines when you configure an IPv6 BGP route reflector:

Because the route reflector forwards routing information between clients, you must make clients of a route reflector fully meshed. If clients are fully meshed, HP recommends disabling route reflection between clients to reduce routing costs.

If a cluster has multiple route reflectors, you must specify the same cluster ID for these route reflectors to avoid routing loops.

To configure an IPv6 BGP route reflector:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Configure the router as a route reflector and specify an IPv6 peer or peer group as a client.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } reflect-client	Not configured by default.
5. Enable route reflection between clients.	reflect between-clients	Optional. Enabled by default.
6. Configure the cluster ID of the route reflector.	reflector cluster-id <i>cluster-id</i>	Optional. By default, a route reflector uses its router ID as the cluster ID.

Configuring BFD for IPv6 BGP

IPv6 BGP maintains neighbor relationships based on the keepalive timer and holdtime timer, which are set in seconds. IPv6 BGP defines that the holdtime interval must be at least three times the keepalive interval. This mechanism makes the detection of a link failure rather slow and thus causes a large quantity of packets to be dropped especially when the failed link is a high-speed link. You can enable BFD to detect the link to a peer. BFD can quickly detect any link failure and thus reduce network convergence time.

After a link failure occurs, BFD may detect the failure before the system performs GR and as a result, GR will fail. Therefore, if GR capability is enabled for IPv6 BGP, use BFD with caution. For more information about BFD, see *High Availability Configuration Guide*.

! IMPORTANT:

Before configuring BFD for IPv6 BGP, you must enable BGP.

To enable BFD for a BGP peer:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable BGP and enter BGP view.	bgp as-number	Not enabled by default.
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Enable BFD for the specified BGP peer.	peer ipv6-address bfd	Not enabled for any BGP peer by default.

Displaying and maintaining IPv6 BGP

Displaying BGP

Task	Command	Remarks
Display IPv6 BGP peer group information.	display bgp ipv6 group [<i>ipv6-group-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 BGP advertised routing information.	display bgp ipv6 network [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 BGP AS path information.	display bgp ipv6 paths [<i>as-regular-expression</i> { begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 BGP peer or peer group information.	display bgp ipv6 peer [<i>group-name</i> log-info <i>ipv4-address</i> verbose <i>ipv6-address</i> { log-info verbose } verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display the prefix entries in the ORF information of the specified BGP peer.	display bgp ipv6 peer { <i>ip-address</i> <i>ipv6-address</i> } received ipv6-prefix [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 BGP routing table information.	display bgp ipv6 routing-table [<i>ipv6-address</i> <i>prefix-length</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 BGP routing information matching an AS path ACL.	display bgp ipv6 routing-table as-path-acl <i>as-path-acl-number</i> [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 BGP routing information with the specified community attribute.	display bgp ipv6 routing-table community [<i>aa:nn</i> <1-13>] [no-advertise no-export no-export-subconfed]* [whole-match] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 BGP routing information matching an IPv6 BGP community list.	display bgp ipv6 routing-table community-list { { <i>basic-community-list-number</i> <i>comm-list-name</i> } [whole-match] <i>adv-community-list-number</i> }&<1-16> [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display dampened IPv6 BGP routing information.	display bgp ipv6 routing-table dampened [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 BGP dampening parameter information.	display bgp ipv6 routing-table dampening parameter [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 BGP routing information originated from different ASs.	display bgp ipv6 routing-table different-origin-as [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 BGP routing flap statistics.	display bgp ipv6 routing-table flap-info [regular-expression <i>as-regular-expression</i> [as-path-acl <i>as-path-acl-number</i> <i>ipv6-address</i> <i>prefix-length</i> [longer-match]] [{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display BGP routing information to or from an IPv4 or IPv6 peer.	display bgp ipv6 routing-table peer { <i>ipv4-address</i> <i>ipv6-address</i> } { advertised-routes received-routes } [<i>network-address</i> <i>prefix-length</i> statistic] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 BGP routing information matching a regular expression.	display bgp ipv6 routing-table regular-expression <i>as-regular-expression</i>	Available in any view
Display IPv6 BGP routing statistics.	display bgp ipv6 routing-table statistic [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Resetting IPv6 BGP connections

Task	Command	Remarks
Perform soft reset on IPv6 BGP connections.	refresh bgp ipv6 { <i>ipv4-address</i> <i>ipv6-address</i> all external group <i>ipv6-group-name</i> internal } { export import }	Available in user view
Reset IPv6 BGP connections.	reset bgp ipv6 { <i>as-number</i> <i>ipv4-address</i> <i>ipv6-address</i> flap-info] all external group <i>group-name</i> internal }	Available in user view

Clearing IPv6 BGP information

Task	Command	Remarks
Clear dampened IPv6 BGP routing information and release suppressed routes.	reset bgp ipv6 dampening [<i>ipv6-address</i> / <i>prefix-length</i>]	Available in user view
Clear IPv6 BGP route flap information.	reset bgp ipv6 flap-info [<i>ipv6-address/prefix-length</i> as-path-acl <i>as-path-acl-number</i> regex <i>as-path-regexp</i>]	Available in user view

IPv6 BGP configuration examples

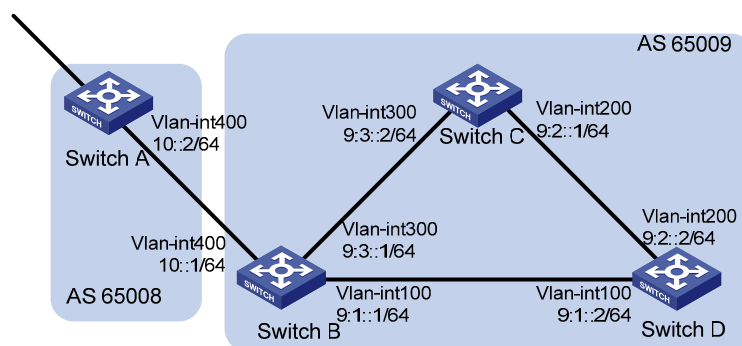
Some examples for IPv6 BGP configuration are similar to those of BGP4. For more information, see "Configuring BGP."

IPv6 BGP basic configuration example

Network requirements

All switches in Figure 115 run IPv6 BGP. Between Switch A and Switch B is an EBGP connection. Switch B, Switch C, and Switch D are fully meshed through IBGP connections.

Figure 115 Network diagram



Configuration procedure

1. Configure IPv6 addresses for interfaces. (Details not shown.)
2. Configure IBGP connections:
Configure Switch B.

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] bgp 65009
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] ipv6-family
[SwitchB-bgp-af-ipv6] peer 9:1::2 as-number 65009
[SwitchB-bgp-af-ipv6] peer 9:3::2 as-number 65009
[SwitchB-bgp-af-ipv6] quit
[SwitchB-bgp] quit
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] bgp 65009
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] ipv6-family
[SwitchC-bgp-af-ipv6] peer 9:3::1 as-number 65009
[SwitchC-bgp-af-ipv6] peer 9:2::2 as-number 65009
[SwitchC-bgp-af-ipv6] quit
[SwitchC-bgp] quit
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] ipv6
[SwitchD] bgp 65009
[SwitchD-bgp] router-id 4.4.4.4
[SwitchD-bgp] ipv6-family
[SwitchD-bgp-af-ipv6] peer 9:1::1 as-number 65009
[SwitchD-bgp-af-ipv6] peer 9:2::1 as-number 65009
[SwitchD-bgp-af-ipv6] quit
[SwitchD-bgp] quit
```

3. Configure the EBGp connection:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] bgp 65008
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] ipv6-family
[SwitchA-bgp-af-ipv6] peer 10::1 as-number 65009
[SwitchA-bgp-af-ipv6] quit
[SwitchA-bgp] quit
```

Configure Switch B.

```
[SwitchB] bgp 65009
[SwitchB-bgp] ipv6-family
[SwitchB-bgp-af-ipv6] peer 10::2 as-number 65008
```

Display IPv6 peer information on Switch B.

```
[SwitchB] display bgp ipv6 peer
```

```

BGP local router ID : 2.2.2.2
Local AS number : 65009
Total number of peers : 3                Peers in established state : 3

```

Peer	AS	MsgRcvd	MsgSent	OutQ	PrefRcv	Up/Down	State
10::2	65008	3	3	0	0	00:01:16	Established
9:3::2	65009	2	3	0	0	00:00:40	Established
9:1::2	65009	2	4	0	0	00:00:19	Established

Display IPv6 peer information on Switch C.

```
[SwitchC] display bgp ipv6 peer
```

```

BGP local router ID : 3.3.3.3
Local AS number : 65009
Total number of peers : 2                Peers in established state : 2

```

Peer	AS	MsgRcvd	MsgSent	OutQ	PrefRcv	Up/Down	State
9:3::1	65009	4	4	0	0	00:02:18	Established
9:2::2	65009	4	5	0	0	00:01:52	Established

Switch A and B have established an EBGP connection; Switch B, C, and D have established IBGP connections with each other.

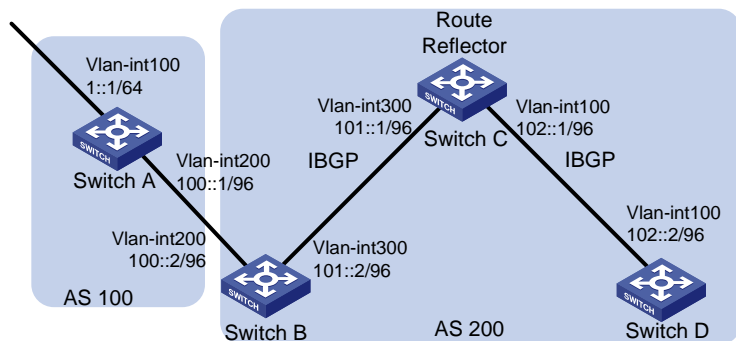
IPv6 BGP route reflector configuration example

Network requirements

As shown in the following figure, Switch B receives an EBGP update and sends it to Switch C, which is configured as a route reflector with two clients: Switch B and Switch D.

Switch B and Switch D do not need to establish an IBGP connection because Switch C reflects updates between them.

Figure 116 Network diagram



Configuration procedure

1. Configure IPv6 addresses for VLAN interfaces. (Details not shown.)

2. Configure IPv6 BGP basic functions:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] bgp 100
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] ipv6-family
[SwitchA-bgp-af-ipv6] peer 100::2 as-number 200
[SwitchA-bgp-af-ipv6] network 1:: 64
```

#Configure Switch B.

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] bgp 200
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] ipv6-family
[SwitchB-bgp-af-ipv6] peer 100::1 as-number 100
[SwitchB-bgp-af-ipv6] peer 101::1 as-number 200
[SwitchB-bgp-af-ipv6] peer 101::1 next-hop-local
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] bgp 200
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] ipv6-family
[SwitchC-bgp-af-ipv6] peer 101::2 as-number 200
[SwitchC-bgp-af-ipv6] peer 102::2 as-number 200
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] ipv6
[SwitchD] bgp 200
[SwitchD-bgp] router-id 4.4.4.4
[SwitchD-bgp] ipv6-family
[SwitchD-bgp-af-ipv6] peer 102::1 as-number 200
```

3. Configure route reflector:

Configure Switch C as a route reflector, and configure Switch B and Switch D as its clients.

```
[SwitchC-bgp-af-ipv6] peer 101::2 reflect-client
[SwitchC-bgp-af-ipv6] peer 102::2 reflect-client
```

Use the **display bgp ipv6 routing-table** command on Switch B and Switch D; both of them learned the network 1::/64.

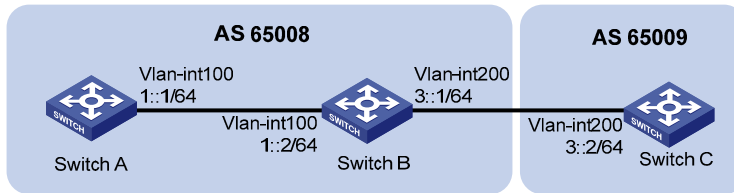
IPv6 BGP IPsec policy configuration example

Network requirements

As shown in the following figure,

- Configure IPv6 BGP on the switches. Switches A and B establish an IBGP relationship. Switches B and C establish an EBGP relationship.
- Configure IPsec policies on the switches to authenticate and encrypt protocol packets.

Figure 117 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure the IBGP connection:

Configure Switch A.

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] bgp 65008
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] ipv6-family
[SwitchA-bgp-af-ipv6] group ibgp internal
[SwitchA-bgp-af-ipv6] peer 1::2 group ibgp
[SwitchA-bgp-af-ipv6] quit
[SwitchA-bgp] quit
```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] bgp 65008
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] ipv6-family
[SwitchB-bgp-af-ipv6] group ibgp internal
[SwitchB-bgp-af-ipv6] peer 1::1 group ibgp
[SwitchB-bgp-af-ipv6] quit
[SwitchB-bgp] quit
```

3. Configure the EBGP connection:

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] ipv6
[SwitchC] bgp 65009
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] ipv6-family
[SwitchC-bgp-af-ipv6] group ebgp external
[SwitchC-bgp-af-ipv6] peer 3::1 as-number 65008
[SwitchC-bgp-af-ipv6] peer 3::1 group ebgp
```



```
[SwitchC-bgp-af-ipv6] quit
[SwitchC-bgp] quit
```

Configure Switch B.

```
[SwitchB-bgp] ipv6-family
[SwitchB-bgp-af-ipv6] group ebgp external
[SwitchB-bgp-af-ipv6] peer 3::2 as-number 65009
[SwitchB-bgp-af-ipv6] peer 3::2 group ebgp
[SwitchB-bgp-af-ipv6] quit
[SwitchB-bgp] quit
```

4. Configure IPsec policies:

On Switch A, create an IPsec proposal named **tran1**, and set the encapsulation mode to transport mode, the security protocol to ESP, the encryption algorithm to DES, and authentication algorithm to SHA1; create an IPsec policy named **policy001**, specify the manual mode for it, reference IPsec proposal **tran1**, set the SPIs of the inbound and outbound SAs to 12345, and the keys for the inbound and outbound SAs using ESP to abcdefg.

```
[SwitchA] ipsec proposal tran1
[SwitchA-ipsec-proposal-tran1] encapsulation-mode transport
[SwitchA-ipsec-proposal-tran1] transform esp
[SwitchA-ipsec-proposal-tran1] esp encryption-algorithm des
[SwitchA-ipsec-proposal-tran1] esp authentication-algorithm sha1
[SwitchA-ipsec-proposal-tran1] quit
[SwitchA] ipsec policy policy001 10 manual
[SwitchA-ipsec-policy-manual-policy001-10] proposal tran1
[SwitchA-ipsec-policy-manual-policy001-10] sa spi outbound esp 12345
[SwitchA-ipsec-policy-manual-policy001-10] sa spi inbound esp 12345
[SwitchA-ipsec-policy-manual-policy001-10] sa string-key outbound esp abcdefg
[SwitchA-ipsec-policy-manual-policy001-10] sa string-key inbound esp abcdefg
[SwitchA-ipsec-policy-manual-policy001-10] quit
```

On Switch B, create an IPsec proposal named **tran1**, and set the encapsulation mode to transport mode, the security protocol to ESP, the encryption algorithm to DES, and authentication algorithm to SHA1; create an IPsec policy named **policy001**, specify the manual mode for it, reference IPsec proposal **tran1**, set the SPIs of the inbound and outbound SAs to 12345, and the keys for the inbound and outbound SAs using ESP to abcdefg; create an IPsec proposal named **tran2**, and set the encapsulation mode to transport mode, the security protocol to ESP, the encryption algorithm to DES, and authentication algorithm to SHA1; create an IPsec policy named **policy002**, specify the manual mode for it, reference IPsec proposal **tran2**, set the SPIs of the inbound and outbound SAs to 54321, and the keys for the inbound and outbound SAs using ESP to gfedcba.

```
[SwitchB] ipsec proposal tran1
[SwitchB-ipsec-proposal-tran1] encapsulation-mode transport
[SwitchB-ipsec-proposal-tran1] transform esp
[SwitchB-ipsec-proposal-tran1] esp encryption-algorithm des
[SwitchB-ipsec-proposal-tran1] esp authentication-algorithm sha1
[SwitchB-ipsec-proposal-tran1] quit
[SwitchB] ipsec policy policy001 10 manual
[SwitchB-ipsec-policy-manual-policy001-10] proposal tran1
```

```

[SwitchB-ipsec-policy-manual-policy001-10] sa spi outbound esp 12345
[SwitchB-ipsec-policy-manual-policy001-10] sa spi inbound esp 12345
[SwitchB-ipsec-policy-manual-policy001-10] sa string-key outbound esp abcdefg
[SwitchB-ipsec-policy-manual-policy001-10] sa string-key inbound esp abcdefg
[SwitchB-ipsec-policy-manual-policy001-10] quit
[SwitchB] ipsec proposal tran2
[SwitchB-ipsec-proposal-tran2] encapsulation-mode transport
[SwitchB-ipsec-proposal-tran2] transform esp
[SwitchB-ipsec-proposal-tran2] esp encryption-algorithm des
[SwitchB-ipsec-proposal-tran2] esp authentication-algorithm sha1
[SwitchB-ipsec-proposal-tran2] quit
[SwitchB] ipsec policy policy002 10 manual
[SwitchB-ipsec-policy-manual-policy002-10] proposal tran2
[SwitchB-ipsec-policy-manual-policy002-10] sa spi outbound esp 54321
[SwitchB-ipsec-policy-manual-policy002-10] sa spi inbound esp 54321
[SwitchB-ipsec-policy-manual-policy002-10] sa string-key outbound esp gfedcba
[SwitchB-ipsec-policy-manual-policy002-10] sa string-key inbound esp gfedcba
[SwitchB-ipsec-policy-manual-policy002-10] quit

```

On Switch C, create an IPsec proposal named **tran2**, and set the encapsulation mode to transport mode, the security protocol to ESP, the encryption algorithm to DES, and authentication algorithm to SHA1; create an IPsec policy named **policy002**, specify the manual mode for it, reference IPsec proposal **tran2**, set the SPIs of the inbound and outbound SAs to 54321, and the keys for the inbound and outbound SAs using ESP to gfedcba.

```

[SwitchC] ipsec proposal tran2
[SwitchC-ipsec-proposal-tran2] encapsulation-mode transport
[SwitchC-ipsec-proposal-tran2] transform esp
[SwitchC-ipsec-proposal-tran2] esp encryption-algorithm des
[SwitchC-ipsec-proposal-tran2] esp authentication-algorithm sha1
[SwitchC-ipsec-proposal-tran2] quit
[SwitchC] ipsec policy policy002 10 manual
[SwitchC-ipsec-policy-manual-policy002-10] proposal tran2
[SwitchC-ipsec-policy-manual-policy002-10] sa spi outbound esp 54321
[SwitchC-ipsec-policy-manual-policy002-10] sa spi inbound esp 54321
[SwitchC-ipsec-policy-manual-policy002-10] sa string-key outbound esp gfedcba
[SwitchC-ipsec-policy-manual-policy002-10] sa string-key inbound esp gfedcba
[SwitchC-ipsec-policy-manual-policy002-10] quit

```

5. Apply IPsec policies to IBGP peers:

Configure Switch A.

```

[SwitchA] bgp 65008
[SwitchA-bgp] ipv6-family
[SwitchA-bgp-af-ipv6] peer 1::2 ipsec-policy policy001
[SwitchA-bgp-af-ipv6] quit
[SwitchA-bgp] quit

```

Configure Switch B.

```

[SwitchB] bgp 65008
[SwitchB-bgp] ipv6-family

```

```
[SwitchB-bgp-af-ipv6] peer 1::1 ipsec-policy policy001
[SwitchB-bgp-af-ipv6] quit
[SwitchB-bgp] quit
```

6. Apply IPsec policies to EBGp peers:

Configure Switch C.

```
[SwitchC] bgp 65009
[SwitchC-bgp] ipv6-family
[SwitchC-bgp-af-ipv6] peer ebgp ipsec-policy policy002
[SwitchC-bgp-af-ipv6] quit
[SwitchC-bgp] quit
```

Configure Switch B.

```
[SwitchB] bgp 65008
[SwitchB-bgp] ipv6-family
[SwitchB-bgp-af-ipv6] peer ebgp ipsec-policy policy002
[SwitchB-bgp-af-ipv6] quit
[SwitchB-bgp] quit
```

7. Verify the configuration:

Display detailed IPv6 BGP peer information.

```
[SwitchB] display bgp ipv6 peer verbose
      BGP Peer is 1::1, remote AS 65008,
      Type: IBGP link
      BGP version 4, remote router ID 1.1.1.1
      BGP current state: Established, Up for 00h01m51s
      BGP current event: RecvKeepalive
      BGP last state: OpenConfirm
      Port: Local - 1029 Remote - 179
      Configured: Active Hold Time: 180 sec Keepalive Time: 60 sec
      Received : Active Hold Time: 180 sec
      Negotiated: Active Hold Time: 180 sec
      Peer optional capabilities:
      Peer support bgp multi-protocol extended
      Peer support bgp route refresh capability
      Address family IPv4 Unicast: advertised and received
Received: Total 0 messages, Update messages 0
Sent: Total 0 messages, Update messages 0
Maximum allowed prefix number: 4294967295
Threshold: 75%
Minimum time between advertisement runs is 30 seconds
Optional capabilities:
Route refresh capability has been enabled
ORF advertise capability based on prefix (type 64):
Local: both
Negotiated: send
Peer Preferred Value: 0
IPsec policy name: policy001, SPI :12345
```

```

Routing policy configured:
No routing policy is configured
    BGP Peer is 3::2, remote AS 65009,
    Type: EBGP link
    BGP version 4, remote router ID 3.3.3.3
    BGP current state: Established, Up for 00h01m51s
    BGP current event: RecvKeepalive
    BGP last state: OpenConfirm
    Port: Local - 1029 Remote - 179
    Configured: Active Hold Time: 180 sec Keepalive Time: 60 sec
    Received : Active Hold Time: 180 sec
    Negotiated: Active Hold Time: 180 sec
    Peer optional capabilities:
    Peer support bgp multi-protocol extended
    Peer support bgp route refresh capability
    Address family IPv4 Unicast: advertised and received

Received: Total 0 messages, Update messages 0
Sent: Total 0 messages, Update messages 0
Maximum allowed prefix number: 4294967295
Threshold: 75%
Minimum time between advertisement runs is 30 seconds
Optional capabilities:
    Route refresh capability has been enabled
    ORF advertise capability based on prefix (type 64):
        Local: both
        Negotiated: send
Peer Preferred Value: 0
IPsec policy name: policy002, SPI :54321
Routing policy configured:
No routing policy is configured

```

The output shows that both IBGP and EBGP neighbor relationships have been established and all protocol packets are protected by IPsec.

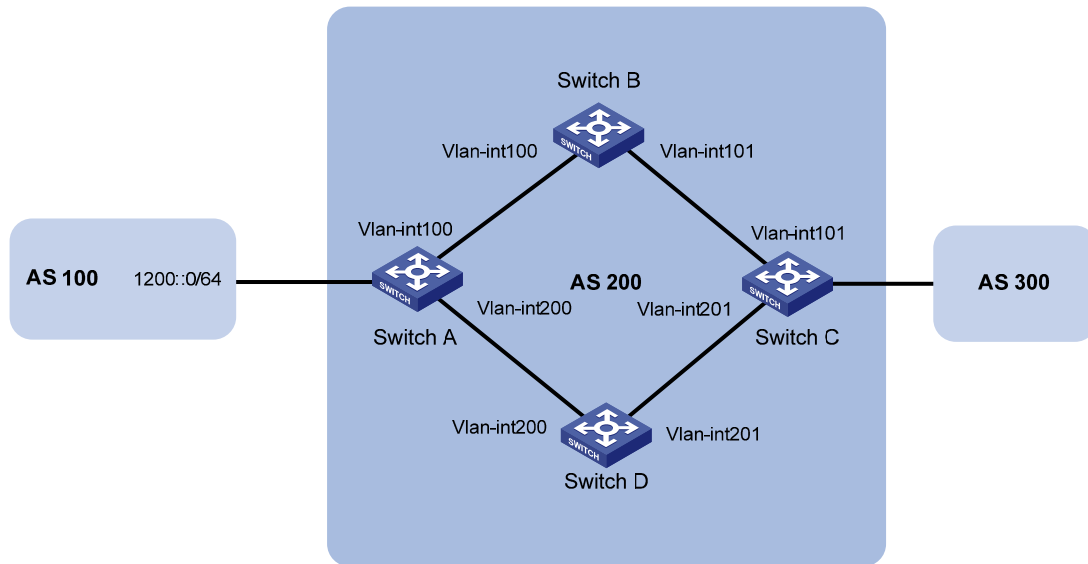
Configuring BFD for IPv6 BGP

Network requirements

As shown in [Figure 118](#),

- Configure OSPFv3 as the IGP in AS 200.
- Establish two IBGP connections between Switch A and Switch C. When both links are working, Switch C adopts the link Switch A \leftrightarrow Switch B \leftrightarrow Switch C to exchange packets with network 1200::0/64. Configure BFD over the link. Then if the link fails, BFD can quickly detect the failure and notify it to IPv6 BGP. Then the link Switch A \leftrightarrow Switch D \leftrightarrow Switch C takes effect immediately.

Figure 118 Network diagram



Device	Interface	IP address	Device	Interface	IP address
Switch A	Vlan-int100	3000::1/64	Switch C	Vlan-int101	3001::3/64
	Vlan-int200	2000::1/64		Vlan-int201	2001::3/64
Switch B	Vlan-int100	3000::2/64	Switch D	Vlan-int200	2000::2/64
	Vlan-int101	3001::2/64		Vlan-int201	2001::2/64

Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure OSPFv3 to make sure that Switch A and Switch C are reachable to each other. (Details not shown.)
3. Configure IPv6 BGP on Switch A.

Establish two IBGP connections between Switch A and Switch C.

```
<SwitchA> system-view
[SwitchA] bgp 200
[SwitchA-bgp] ipv6-family
[SwitchA-bgp-af-ipv6] peer 3001::3 as-number 200
[SwitchA-bgp-af-ipv6] peer 2001::3 as-number 200
[SwitchA-bgp-af-ipv6] quit
```

When the two links between Switch A and Switch C are both up, Switch C adopts the link Switch A<—>Switch B<—>Switch C to exchange packets with network 1200::0/64. (Set a higher MED value for route 1200::0/64 sent to peer 2001::3 on Switch A.)

- o Create IPv6 ACL 2000 to permit 1200::0/64 to pass.


```
[SwitchA] acl ipv6 number 2000
[SwitchA-acl6-basic-2000] rule permit source 1200::0 64
[SwitchA-acl6-basic-2000] quit
```
- o Create two route policies, **apply_med_50** and **apply_med_100**. Policy **apply_med_50** sets the MED for route 1200::0/64 to 50. Policy **apply_med_100** sets that to 100.

```
[SwitchA] route-policy apply_med_50 permit node 10
[SwitchA-route-policy] if-match ipv6 address acl 2000
[SwitchA-route-policy] apply cost 50
[SwitchA-route-policy] quit
[SwitchA] route-policy apply_med_100 permit node 10
[SwitchA-route-policy] if-match ipv6 address acl 2000
[SwitchA-route-policy] apply cost 100
[SwitchA-route-policy] quit
```

- Apply routing policy **apply_med_50** to routes outgoing to peer 3001::3, and apply routing policy **apply_med_100** to routes outgoing to peer 2001::3.

```
[SwitchA] bgp 200
[SwitchA-bgp] ipv6-family
[SwitchA-bgp-af-ipv6] network 1200:: 64
[SwitchA-bgp-af-ipv6] peer 3001::3 route-policy apply_med_50 export
[SwitchA-bgp-af-ipv6] peer 2001::3 route-policy apply_med_100 export
```

Configure BFD over the link to peer 3001::3 so that when the link Switch A<—>Switch B<—>Switch C fails, BFD can quickly detect the failure and notify it to IPv6 BGP, and then the link Switch A<—>Switch D<—>Switch C takes effect immediately.

```
[SwitchA-bgp-af-ipv6] peer 3001::3 bfd
[SwitchA-bgp-af-ipv6] quit
[SwitchA-bgp] quit
```

4. Configure IPv6 BGP on Switch C:

```
<SwitchC> system-view
[SwitchC] bgp 200
[SwitchC-bgp] ipv6-family
[SwitchC-bgp-af-ipv6] peer 3000::1 as-number 200
[SwitchC-bgp-af-ipv6] peer 3000::1 bfd
[SwitchC-bgp-af-ipv6] peer 2000::1 as-number 200
[SwitchC-bgp-af-ipv6] quit
[SwitchC-bgp] quit
```

5. Configure BFD parameters (you can use default BFD parameters instead):

Configure Switch A.

```
[SwitchA] bfd session init-mode active
[SwitchA] interface vlan-interface 100
```

- Configure the minimum interval for transmitting BFD control packets as 500 milliseconds.


```
[SwitchA-Vlan-interface100] bfd min-transmit-interval 500
```
- Configure the minimum interval for receiving BFD control packets as 500 milliseconds.


```
[SwitchA-Vlan-interface100] bfd min-receive-interval 500
```
- Configure the detect multiplier as 7.


```
[SwitchA-Vlan-interface100] bfd detect-multiplier 7
[SwitchA-Vlan-interface100] quit
```

Configure Switch C.

```
[SwitchC] bfd session init-mode active
[SwitchC] interface vlan-interface 101
```

- Configure the minimum interval for transmitting BFD control packets as 500 milliseconds.
[SwitchC-Vlan-interface101] bfd min-transmit-interval 500
- Configure the minimum interval for receiving BFD control packets as 500 milliseconds.
[SwitchC-Vlan-interface101] bfd min-receive-interval 500
- Configure the detect multiplier as 7.
[SwitchC-Vlan-interface101] bfd detect-multiplier 7
[SwitchC-Vlan-interface101] return

6. Verify the configuration:

The following operations are made on Switch C. Operations on Switch A and Switch B are similar and are not shown.

Display detailed BFD session information.

```
<SwitchC> display bfd session verbose
```

```
Total session number: 1   Up session number: 1   Init mode: Active
```

```
IPv6 Session working under Ctrl mode:
```

```

Local Discr: 17                               Remote Discr: 13
Source IP: 3001::3
Destination IP: 3000::1
Session State: Up                             Interface: Vlan-interface101
Min Trans Inter: 500ms                       Act Trans Inter: 500ms
Min Recv Inter: 500ms                       Act Detect Inter: 3000ms
Recv Pkt Num: 57                             Send Pkt Num: 53
Hold Time: 2200ms                           Connect Type: Direct
Running Up for: 00:00:06                     Auth mode: none
Protocol: BGP6
Diag Info: No Diagnostic

```

The output shows that a BFD session is established between Switch A's VLAN-interface 100 and Switch C's VLAN-interface 101 and that BFD runs properly.

Display IPv6 peer information on Switch C, and you can see that the neighborhood between Switch A and Switch C is established.

```
<SwitchC> display bgp ipv6 peer
```

```
BGP local router ID : 1.1.1.1
```

```
Local AS number : 200
```

```
Total number of peers : 2
```

```
Peers in established state : 2
```

Peer	AS	MsgRcvd	MsgSent	OutQ	PrefRcv	Up/Down	State
2000::1	200	7	10	0	0	00:01:05	Established
3000::1	200	7	10	0	0	00:01:34	Established

Display route 1200::0/64 on Switch C, and you can see that Switch A and Switch C communicate through Switch B.

```
<SwitchC> display ipv6 routing-table 1200::0 64 verbose
```

```
Routing Table :
```

```
Summary Count : 2
```

```

Destination : 1200::
NextHop      : 3000::1
RelayNextHop : 3001::2
Neighbor     : 3000::1
Interface    : Vlan-interface101
State        : Active Adv
Tunnel ID    : 0x0
Age          : 4538sec
PrefixLength : 64
Preference   : 255
Tag          : 0H
ProcessID    : 0
Protocol     : BGP4+
Cost         : 50
Label        : NULL

```

```

Destination : 1200::
NextHop      : 2000::1
RelayNextHop : 2001::2
Neighbor     : 2000::1
Interface    : Vlan-interface201
State        : Invalid Adv
Tunnel ID    : 0x0
Age          : 4515sec
PrefixLength : 64
Preference   : 255
Tag          : 0H
ProcessID    : 0
Protocol     : BGP4+
Cost         : 100
Label        : NULL

```

The output shows that Switch C has two routes to reach network 1200::0/64: Switch C<—>Switch B<—>Switch A, which is the currently active route; Switch C<—>Switch D<—>Switch A, which is the backup route.

Enable BFD debugging on Switch C.

```

<SwitchC> debugging bfd scm
<SwitchC> debugging bfd event
<SwitchC> debugging bgp bfd
<SwitchC> terminal monitor
<SwitchC> terminal debugging

```

The following debugging information shows that Switch C can quickly detect the failure on Switch B.

```

%Nov 5 11:42:24:172 2009 SwitchC BFD/5/BFD_CHANGE_FSM: Sess[3001::3/3000::1,
13/17,VLAN101,Ctrl], Sta: UP->DOWN, Diag: 1
%Nov 5 11:42:24:172 2009 SwitchC BGP/5/BGP_STATE_CHANGED: 3000::1 state is changed
from ESTABLISHED to IDLE.
*Nov 5 11:42:24:187 2009 SwitchC RM/6/RMDEBUG: BGP_BFD: Recv BFD DOWN msg, Src IP
3001::3, Dst IP 3000::1, Instance ID 0.
*Nov 5 11:42:24:187 2009 SwitchC RM/6/RMDEBUG: BGP_BFD: Reset BGP session 3000::1
for BFD session down.
*Nov 5 11:42:24:187 2009 SwitchC RM/6/RMDEBUG: BGP_BFD: Send DELETE msg to BFD,
Connection type DIRECT, Src IP 3001::3, Dst IP 3000::1, Instance ID 0.

```

Display route 1200::0/64 on Switch C, and you can see that Switch A and Switch C communicate through Switch D.

```

<SwitchC> display ipv6 routing-table 1200::0 64 verbose
Routing Table :
Summary Count : 1
Destination : 1200::
NextHop      : 2000::1
RelayNextHop : 2001::2
PrefixLength : 64
Preference   : 255
Tag          : 0H

```


Neighbor	: 2000::1	ProcessID	: 0
Interface	: Vlan-interface201	Protocol	: BGP4+
State	: Active Adv	Cost	: 100
Tunnel ID	: 0x0	Label	: NULL
Age	: 4635sec		

The output shows that Switch C has one route to reach network 1200::0/64, that is, Switch C<—>Switch D<—>Switch A.

Troubleshooting IPv6 BGP configuration

IPv6 BGP peer relationship not established

Symptom

Display BGP peer information by using the **display bgp ipv6 peer** command. The state of the connection to the peer cannot become established.

Analysis

To become IPv6 BGP peers, any two routers must establish a TCP session using port 179 and exchange open messages successfully.

Processing steps

1. Use the **display current-configuration** command to check that the peer's AS number is correct.
2. Use the **display bgp ipv6 peer** command to check that the peer's IPv6 address is correct.
3. If a loopback interface is used, check that the loopback interface is specified with the **peer connect-interface** command.
4. If the peer is not directly connected, check that the **peer ebgp-max-hop** command is configured.
5. Check that a valid route to the peer is available.
6. Use the **ping** command to check the connectivity to the peer.
7. Use the **display tcp ipv6 status** command to check the TCP connection.
8. Check whether an ACL for disabling TCP port 179 is configured.

Configuring routing policies

Hardware compatibility

The HP 5500 SI Switch Series does not support OSPF, BGP, IS-IS, OSPFv3, IPv6 BGP, IPv6 IS-IS, or FRR.

Introduction to routing policy

Routing policies are used to receive, advertise, and redistribute only specific routes and modify the attributes of some routes.

Routing policy in this chapter involves both IPv4 routing policy and IPv6 routing policy.

A routing policy is used to filter routes when they are received, advertised, or redistributed and modify the attributes of some routes.

Routing policy application

A routing policy has the following applications:

- Filters advertised routes.
- Filters received routes.
- Filters redistributed routes.
- Modifies or sets the attributes of some routes.

Routing policy implementation

To configure a routing policy, you must do the following:

1. Define some filters based on the attributes of routing information, such as destination address, and the advertising router's address.
2. Apply the filters to the routing policy.

You can use multiple filters to define match criteria. For detailed information, see "[Filters](#)."

Filters

You can use the following types of filters: ACL, IP prefix list, AS path ACL, community list, extended community list, and routing policy.

ACL

ACL involves IPv4 ACL and IPv6 ACL. An ACL is configured to match the destinations or next hops of routing information.

For more information about ACL, see *ACL and QoS Configuration Guide*.

IP prefix list

IP prefix list involves IPv4 prefix list and IPv6 prefix list.

An IP prefix list is configured to match the destination address of routing information. You can use the **gateway** option to allow only routing information from certain routers to be received. For **gateway** option information, see "[Configuring RIP](#)" and "[Configuring OSPF](#)."

An IP prefix list, identified by name, can comprise multiple items. Each item, identified by an index number, can specify a prefix range to match. An item with a smaller index number is matched first. If one item is matched, the IP prefix list is passed, and the packet will not go to the next item.

AS-PATH list

An AS-PATH list, configured based on the BGP AS PATH attribute, can only be used to match BGP routing information.

For more information about AS-PATH list, see "[Configuring BGP](#)."

Community list

A community list, configured based on the BGP community attribute, can only be used to match BGP routing information.

For more information about community list, see "[Configuring BGP](#)."

Extended community list

An extended community list, configured based on the BGP extended community attribute (Route-Target for VPN and Source of Origin), can only be used to match BGP routing information.

Routing policy

A routing policy is used to match routing information and modify the attributes of permitted routes. It can reference the filters to define its own match criteria.

A routing policy can comprise multiple nodes, which are in logic OR relationship. Each routing policy node is a match unit, and a node with a smaller number is matched first. Once a node is matched, the routing policy is passed and the packet will not go to the next node.

A routing policy node comprises a set of **if-match**, **apply**, and **continue** clauses.

- The **if-match** clauses define the match criteria. The matching objects are some attributes of routing information. The **if-match** clauses of a routing policy node is in a logical AND relationship. A packet must match all the **if-match** clauses of the node to pass it.
- The **apply** clauses of the node specify the actions to be taken on the permitted packets, such as modifying a route attribute.
- The **continue** clause specifies the next routing policy node to be matched. With this clause configured, when a route matches the current routing policy node, it continues to match against the specified next node in the same routing policy. The **continue** clause combines the **if-match** and **apply** clauses of the two nodes to improve flexibility of the routing policy.

Follow these guidelines when you configure **if-match**, **apply**, and **continue** clauses:

- If you want to implement route filtering only, you do not need to configure **apply** clauses.
- If you do not configure any **if-match** clauses for a permit-mode node, the node permits all routes to pass.
- Configure a permit-mode node containing no **if-match** or **apply** clauses behind multiple deny-mode nodes to allow unmatched routes to pass.

Defining filters

Configuration prerequisites

Before you configure this task, you must determine IP-prefix list name, matching address range, and extcommunity list sequence number.

Defining an IP-prefix list

Defining an IPv4 prefix list

Identified by name, an IPv4 prefix list can comprise multiple items. Each item specifies a prefix range to match and is identified by an index number.

An item with a smaller index number is matched first. If one item is matched, the IP prefix list is passed, and the routing information will not go to the next item.

To define an IPv4 prefix list:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Define an IPv4 prefix list.	ip ip-prefix <i>ip-prefix-name</i> [index <i>index-number</i>] { deny permit } <i>ip-address mask-length</i> [greater-equal <i>min-mask-length</i>] [less-equal <i>max-mask-length</i>]	Not defined by default.

If all the items are set to the **deny** mode, no routes can pass the IPv4 prefix list. You must define the **permit** 0.0.0.0 0 **less-equal** 32 item following multiple **deny** items to allow other IPv4 routing information to pass.

For example, the following configuration filters routes 10.1.0.0/16, 10.2.0.0/16, and 10.3.0.0/16, but allows other routes to pass.

```
<Sysname> system-view
[Sysname] ip ip-prefix abc index 10 deny 10.1.0.0 16
[Sysname] ip ip-prefix abc index 20 deny 10.2.0.0 16
[Sysname] ip ip-prefix abc index 30 deny 10.3.0.0 16
[Sysname] ip ip-prefix abc index 40 permit 0.0.0.0 0 less-equal 32
```

Defining an IPv6 prefix list

Identified by name, each IPv6 prefix list can comprise multiple items. Each item specifies a prefix range to match and is identified by an index number.

An item with a smaller index number is matched first. If one item is matched, the IPv6 prefix list is passed, and the routing information will not go to the next item.

To define an IPv6 prefix list:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Define an IPv6 prefix list.	ip ipv6-prefix <i>ipv6-prefix-name</i> [index <i>index-number</i>] { deny permit } <i>ipv6-address prefix-length</i> [greater-equal <i>min-prefix-length</i>] [less-equal <i>max-prefix-length</i>]	Not defined by default.

If all items are set to the **deny** mode, no routes can pass the IPv6 prefix list. You must define the **permit** :: 0 **less-equal** 128 item following multiple **deny** items to allow other IPv6 routing information to pass.

For example, the following configuration filters routes 2000:1::/48, 2000:2::/48, and 2000:3::/48, but allows other routes to pass.

```
<Sysname> system-view
[Sysname] ip ipv6-prefix abc index 10 deny 2000:1:: 48
[Sysname] ip ipv6-prefix abc index 20 deny 2000:2:: 48
[Sysname] ip ipv6-prefix abc index 30 deny 2000:3:: 16
[Sysname] ip ipv6-prefix abc index 40 permit :: 0 less-equal 128
```

Defining an AS path list

You can define multiple items for an AS path list that is identified by number. The relationship between items is logical OR. If a route matches one of these items, it passes the AS path list.

To define an AS path list:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Define an AS path ACL.	ip as-path <i>as-path-number</i> { deny permit } <i>regular-expression</i>	Not defined by default.

Defining a community list

You can define multiple items for a community list that is identified by number. During matching, the relationship between items is logic OR. If routing information matches one of these items, it passes the community list.

To define a community list:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Define a community list.	<ul style="list-style-type: none"> Define a basic community list: ip community-list { <i>basic-comm-list-num</i> basic <i>comm-list-name</i> } { deny permit } [<i>community-number-list</i>] [internet no-advertise no-export no-export-subconfed] * Define an advanced community list: ip community-list { <i>adv-comm-list-num</i> advanced <i>comm-list-name</i> } { deny permit } <i>regular-expression</i> 	Use either approach. Not defined by default.

Defining an extended community list

You can define multiple items for an extended community list that is identified by number. During matching, the relation between items is logic OR. If routing information matches one of these items, it passes the extended community list.

To define an extended community list:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Define an extended community list.	ip extcommunity-list <i>ext-comm-list-number</i> { deny permit } { rt <i>route-target</i> }&<1-16>	Not defined by default.

Configuring a routing policy

A routing policy is used to filter routing information and modify attributes of matching routing information. The match criteria of a routing policy can be configured by referencing the appropriate filters.

A routing policy can comprise multiple nodes, and each routing policy node contains the following:

- **if-match clauses**—Define the match criteria that routing information must satisfy. The matching objects are some attributes of routing information.
- **apply clauses**—Specify the actions to be taken on routing information that has satisfied the match criteria, such as route attribute modification.
- **continue clauses**—Specify the next routing policy node to be matched. With this clause configured, when a route matches the current routing policy node, it continues to match against the specified next node in the same routing policy.

Configuration prerequisites

Before configuring this task, you need to configure the filters and routing protocols. You also need to decide on the name of the routing policy and node numbers, match criteria, and attributes to be modified.

Creating a routing policy

Follow these guidelines when you create a routing policy:

- If a routing policy node has the **permit** keyword specified, routing information matching all the **if-match** clauses of the node will be handled using the **apply** clauses of this node, without needing to match against the next node. If routing information does not match the node, it goes to the next node for a match.
- If a routing policy node has the **deny** keyword specified, the **apply** clauses of the node is not executed. When routing information matches all the **if-match** clauses of the node, it cannot pass the node, or go to the next node. If route information cannot match all the **if-match** clauses of the node, it goes to the next node for a match.
- When a routing policy has more than one node, at least one node must be configured with the **permit** keyword. If the routing policy is used to filter routing information, routing information that does not meet any node cannot pass the routing policy. If all nodes of the routing policy are set with the **deny** keyword, no routing information can pass it.

To create a routing policy:

Step	Command
1. Enter system view.	system-view
2. Create a routing policy, specify a node for it, and enter routing policy view.	route-policy <i>route-policy-name</i> { deny permit } node <i>node-number</i>

Defining if-match clauses

Follow these guidelines when you define **if-match** clauses:

- The **if-match** clauses of a routing policy node are in logic AND relationship. Routing information must satisfy all of its **if-match** clauses before being executed with its **apply** clauses. If an **if-match** command exceeds the maximum length, multiple identical **if-match** clauses are generated. These clauses are in logical OR relationship. Routing information only needs to match one of them.
- You can specify any number of **if-match** clauses for a routing policy node. If no **if-match** clause is specified, and the routing policy node is in **permit** mode, all routing information can pass the node. If it is in **deny** mode, no routing information can pass it.
- If the ACL referenced by an **if-match** clause does not exist, the clause is always satisfied; if no rules of the referenced ACL are matched or the matching rule is inactive, the clause is not satisfied.
- An ACL specified in an **if-match** clause must be a non-VPN ACL.
- The **if-match** commands for matching IPv4 destination, next hop, and source address are different from those for matching IPv6 ones.
- BGP does not support criteria for matching against outbound interfaces of routing information.

To define **if-match** clauses:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter routing policy view.	route-policy <i>route-policy-name</i> { deny permit } node <i>node-number</i>	N/A
3. Define match criteria for IPv4 routes.	<ul style="list-style-type: none"> • Match IPv4 routing information specified in the ACL: if-match acl <i>acl-number</i> • Match IPv4 routing information specified in the IP prefix list: if-match ip-prefix <i>ip-prefix-name</i>. • Match IPv4 routing information whose next hop or source is specified in the ACL or IP prefix list: if-match ip { next-hop route-source } { acl <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } 	Optional. Not configured by default.
4. Match IPv6 routing information whose next hop or source is specified in the ACL or IP prefix list.	if-match ipv6 { address next-hop route-source } { acl <i>acl-number</i> prefix-list <i>ipv6-prefix-name</i> }	Optional. Not configured by default.

Step	Command	Remarks
5. Match BGP routing information whose AS path attribute is specified in the AS path lists.	if-match as-path <i>AS-PATH-number</i> &<1-16>	Optional. Not configured by default.
6. Match BGP routing information whose community attribute is specified in the community lists.	if-match community { { <i>basic-community-list-number</i> <i>comm-list-name</i> } [whole-match] <i>adv-community-list-number</i> }&<1-16>	Optional. Not configured by default.
7. Match routes having the specified cost.	if-match cost <i>value</i>	Optional. Not configured by default.
8. Match BGP routing information whose extended community attribute is specified in the extended community list(s).	if-match extcommunity <i>ext-comm-list-number</i> &<1-16>	Optional. Not configured by default.
9. Match routing information having specified outbound interface(s).	if-match interface { <i>interface-type</i> <i>interface-number</i> }&<1-16>	Optional. Not configured by default.
10. Match routing information having the specified route type.	if-match route-type { external-type1 external-type1or2 external-type2 internal is-is-level-1 is-is-level-2 nssa-external-type1 nssa-external-type1or2 nssa-external-type2 } *	Optional. Not configured by default.
11. Match RIP, OSPF, and IS-IS routing information having the specified tag value.	if-match tag <i>value</i>	Optional. Not configured by default.

Defining apply clauses

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter routing policy view.	route-policy <i>route-policy-name</i> { deny permit } node <i>node-number</i>	Not created by default.
3. Set the AS-PATH attribute for BGP routing information.	apply as-path <i>as-number</i> &<1-10> [replace]	Optional. Not set by default.
4. Delete the community attribute of BGP routing information using the community list.	apply comm-list { <i>comm-list-number</i> <i>comm-list-name</i> } delete	Optional. Not configured by default.
5. Set the community attribute for BGP routing information.	apply community { none additive { <i>community-number</i> &<1-16> <i>aa:nn</i> &<1-16> } internet no-advertise no-export no-export-subconfed } * [additive] }	Optional. Not set by default.

Step	Command	Remarks
6. Set a cost for routing information.	apply cost [+ -] <i>value</i>	Optional. Not set by default.
7. Set a cost type for routing information.	apply cost-type [external internal type-1 type-2]	Optional. Not set by default.
8. Set the extended community attribute for BGP routing.	apply extcommunity { { rt route-target } &<1-16> [additive] }	Optional. Not set by default.
9. Set the next hop.	<ul style="list-style-type: none"> Set the next hop for IPv4 routes: apply ip-address next-hop <i>ip-address</i> Set the next hop for IPv6 routes: apply ipv6 next-hop <i>ipv6-address</i> 	Optional. Not set by default. The setting does not apply to redistributed routing information. The apply ip-address next-hop and apply ipv6 next-hop commands do not apply to redistributed IPv4 and IPv6 routes.
10. Inject routing information to a specified ISIS level.	apply isis { level-1 level-1-2 level-2 }	Optional. Not configured by default.
11. Set the local preference for BGP routing information.	apply local-preference <i>preference</i>	Optional. Not set by default.
12. Set the origin attribute for BGP routing information.	apply origin { egp <i>as-number</i> igp incomplete }	Optional. Not set by default.
13. Set the preference for the routing protocol.	apply preference <i>preference</i>	Optional. Not set by default.
14. Set a preferred value for BGP routing information.	apply preferred-value <i>preferred-value</i>	Optional. Not set by default.
15. Set a tag value for RIP, OSPF, or IS-IS routing information.	apply tag <i>value</i>	Optional. Not set by default.
16. Configure FRR.	apply fast-reroute { backup-interface <i>interface-type interface-number</i> [backup-nexthop <i>ip-address</i>] }	Optional. Not configured by default.

NOTE:

The difference between IPv4 and IPv6 **apply** clauses is the command for setting the next hop for routing information.

Defining a continue clause

Follow these guidelines when you define a **continue** clause:

- If you configure the same type of **apply** clauses that set different values (including the **apply community** and **apply extcommunity** clauses with the **additive** keyword) on nodes that are

combined by the **continue** clause, the **apply** clause configured on the last matching node takes effect.

- If you configure the **apply community** clause for multiple nodes that are combined by the **continue** clause, the **apply comm-list delete** clause configured on the current node cannot delete the community attributes of preceding nodes.

To define a **continue** clause:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a routing policy and enter routing policy view.	route-policy <i>route-policy-name</i> { deny permit } node <i>node-number</i>	Not created by default.
3. Specify the next routing policy node to be matched.	continue [<i>node-number</i>]	Optional. Not configured by default. The node number specified must be larger than the current node number.

Displaying and maintaining the routing policy

Task	Command	Remarks
Display BGP AS-PATH list information.	display ip as-path [<i>as-path-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display BGP community list information.	display ip community-list [<i>basic-community-list-number</i> <i>adv-community-list-number</i> <i>comm-list-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display BGP extended community list information.	display ip extcommunity-list [<i>ext-comm-list-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv4 prefix list statistics.	display ip ip-prefix [<i>ip-prefix-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 prefix list statistics.	display ip ipv6-prefix [<i>ipv6-prefix-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display routing policy information.	display route-policy [<i>route-policy-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear IPv4 prefix list statistics.	reset ip ip-prefix [<i>ip-prefix-name</i>]	Available in user view
Clear IPv6 prefix list statistics.	reset ip ipv6-prefix [<i>ipv6-prefix-name</i>]	Available in user view

Routing policy configuration examples

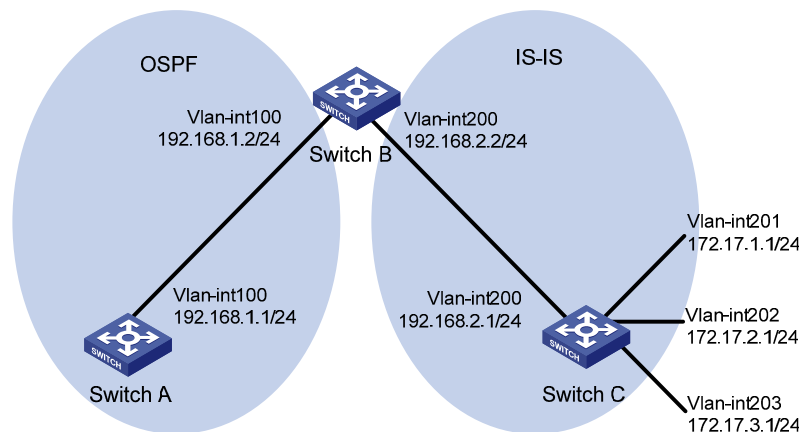
Applying a routing policy to IPv4 route redistribution

Network requirements

As shown in Figure 119, Switch B exchanges routing information with Switch A using OSPF, and with Switch C using IS-IS.

On Switch B, enable route redistribution from IS-IS to OSPF, and apply a routing policy to set the cost of route 172.17.1.0/24 to 100 and the tag of route 172.17.2.0/24 to 20.

Figure 119 Network diagram



Configuration procedure

1. Specify IP addresses for interfaces. (Details not shown.)
2. Configure IS-IS:

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] isis
[SwitchC-isis-1] is-level level-2
[SwitchC-isis-1] network-entity 10.0000.0000.0001.00
[SwitchC-isis-1] quit
[SwitchC] interface vlan-interface 200
[SwitchC-vlan-interface200] isis enable
[SwitchC-vlan-interface200] quit
[SwitchC] interface vlan-interface 201
[SwitchC-vlan-interface201] isis enable
[SwitchC-vlan-interface201] quit
[SwitchC] interface vlan-interface 202
[SwitchC-vlan-interface202] isis enable
[SwitchC-vlan-interface202] quit
[SwitchC] interface vlan-interface 203
[SwitchC-vlan-interface203] isis enable
[SwitchC-vlan-interface203] quit
```

Configure Switch B.

```

<SwitchB> system-view
[SwitchB] isis
[SwitchB-isis-1] is-level level-2
[SwitchB-isis-1] network-entity 10.0000.0000.0002.00
[SwitchB-isis-1] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] isis enable
[SwitchB-Vlan-interface200] quit

```

3. Configure OSPF and route redistribution:

Configure OSPF on Switch A.

```

<SwitchA> system-view
[SwitchA] ospf
[SwitchA-ospf-1] area 0
[SwitchA-ospf-1-area-0.0.0.0] network 192.168.1.0 0.0.0.255
[SwitchA-ospf-1-area-0.0.0.0] quit
[SwitchA-ospf-1] quit

```

On Switch B, configure OSPF and enable route redistribution from IS-IS.

```

[SwitchB] ospf
[SwitchB-ospf-1] area 0
[SwitchB-ospf-1-area-0.0.0.0] network 192.168.1.0 0.0.0.255
[SwitchB-ospf-1-area-0.0.0.0] quit
[SwitchB-ospf-1] import-route isis 1
[SwitchB-ospf-1] quit

```

Display the OSPF routing table on Switch A to view redistributed routes.

```

[SwitchA] display ospf routing

```

```

          OSPF Process 1 with Router ID 192.168.1.1
          Routing Tables

```

```

Routing for Network

```

Destination	Cost	Type	NextHop	AdvRouter	Area
192.168.1.0/24	1562	Stub	192.168.1.1	192.168.1.1	0.0.0.0

```

Routing for ASEs

```

Destination	Cost	Type	Tag	NextHop	AdvRouter
172.17.1.0/24	1	Type2	1	192.168.1.2	192.168.2.2
172.17.2.0/24	1	Type2	1	192.168.1.2	192.168.2.2
172.17.3.0/24	1	Type2	1	192.168.1.2	192.168.2.2
192.168.2.0/24	1	Type2	1	192.168.1.2	192.168.2.2

```

Total Nets: 5

```

```

Intra Area: 1 Inter Area: 0 ASE: 4 NSSA: 0

```

4. Configure filtering lists:

Configure ACL 2002 to permit route 172.17.2.0/24.

```

[SwitchB] acl number 2002
[SwitchB-acl-basic-2002] rule permit source 172.17.2.0 0.0.0.255
[SwitchB-acl-basic-2002] quit

```

```
# Configure IP prefix list prefix-a to permit route 172.17.1.0/24.
[SwitchB] ip ip-prefix prefix-a index 10 permit 172.17.1.0 24
```

5. Configure a routing policy:

```
[SwitchB] route-policy isis2ospf permit node 10
[SwitchB-route-policy] if-match ip-prefix prefix-a
[SwitchB-route-policy] apply cost 100
[SwitchB-route-policy] quit
[SwitchB] route-policy isis2ospf permit node 20
[SwitchB-route-policy] if-match acl 2002
[SwitchB-route-policy] apply tag 20
[SwitchB-route-policy] quit
[SwitchB] route-policy isis2ospf permit node 30
[SwitchB-route-policy] quit
```

6. Apply the routing policy to route redistribution:

On Switch B, apply the routing policy when redistributing routes.

```
[SwitchB] ospf
[SwitchB-ospf-1] import-route isis 1 route-policy isis2ospf
[SwitchB-ospf-1] quit
```

Display the OSPF routing table on Switch A. The cost of route 172.17.1.0/24 is 100, the tag of route 172.17.1.0/24 is 20.

```
[SwitchA] display ospf routing
```

```
OSPF Process 1 with Router ID 192.168.1.1
Routing Tables
```

Routing for Network

Destination	Cost	Type	NextHop	AdvRouter	Area
192.168.1.0/24	1	Transit	192.168.1.1	192.168.1.1	0.0.0.0

Routing for ASEs

Destination	Cost	Type	Tag	NextHop	AdvRouter
172.17.1.0/24	100	Type2	1	192.168.1.2	192.168.2.2
172.17.2.0/24	1	Type2	20	192.168.1.2	192.168.2.2
172.17.3.0/24	1	Type2	1	192.168.1.2	192.168.2.2
192.168.2.0/24	1	Type2	1	192.168.1.2	192.168.2.2

Total Nets: 5

Intra Area: 1 Inter Area: 0 ASE: 4 NSSA: 0

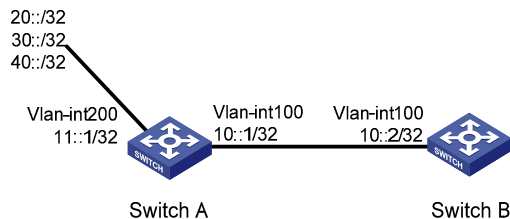
Applying a routing policy to IPv6 route redistribution

Network requirements

As shown in [Figure 120](#):

- Enable RIPng on Switch A and Switch B.
- On Switch A, configure three static routes, and apply a routing policy to static route redistribution to permit routes 20::0/32 and 40::0/32, and deny route 30::0/32.
- Display RIPng routing table information on Switch B to verify the configuration.

Figure 120 Network diagram



Configuration procedure

1. Configure Switch A:

Configure IPv6 addresses for VLAN-interface 100 and VLAN-interface 200.

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ipv6 address 10::1 32
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 200
[SwitchA-Vlan-interface200] ipv6 address 11::1 32
[SwitchA-Vlan-interface200] quit
```

Enable RIPng on VLAN-interface 100.

```
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ripng 1 enable
[SwitchA-Vlan-interface100] quit
```

Configure three static routes—with next hop 11::2—and make sure that they are active.

```
[SwitchA] ipv6 route-static 20:: 32 11::2
[SwitchA] ipv6 route-static 30:: 32 11::2
[SwitchA] ipv6 route-static 40:: 32 11::2
```

Configure a routing policy.

```
[SwitchA] ip ipv6-prefix a index 10 permit 30:: 32
[SwitchA] route-policy static2ripng deny node 0
[SwitchA-route-policy] if-match ipv6 address prefix-list a
[SwitchA-route-policy] quit
[SwitchA] route-policy static2ripng permit node 10
[SwitchA-route-policy] quit
```

Enable RIPng and apply the routing policy to static route redistribution.

```
[SwitchA] ripng
[SwitchA-ripng-1] import-route static route-policy static2ripng
```

2. Configure Switch B:

Configure the IPv6 address for VLAN-interface 100.

```
[SwitchB] ipv6
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ipv6 address 10::2 32
```

Enable RIPng on VLAN-interface 100.

```
[SwitchB-Vlan-interface100] ripng 1 enable
[SwitchB-Vlan-interface100] quit
```

```

# Enable RIPng.
[SwitchB] ripng
# Display RIPng routing table information.
[SwitchB-ripng-1] display ripng 1 route
    Route Flags: A - Aging, S - Suppressed, G - Garbage-collect
    -----

Peer FE80::7D58:0:CA03:1 on Vlan-interface 100
Dest 10::/32,
    via FE80::7D58:0:CA03:1, cost 1, tag 0, A, 18 Sec
Dest 20::/32,
    via FE80::7D58:0:CA03:1, cost 1, tag 0, A, 8 Sec
Dest 40::/32,
    via FE80::7D58:0:CA03:1, cost 1, tag 0, A, 3 Sec

```

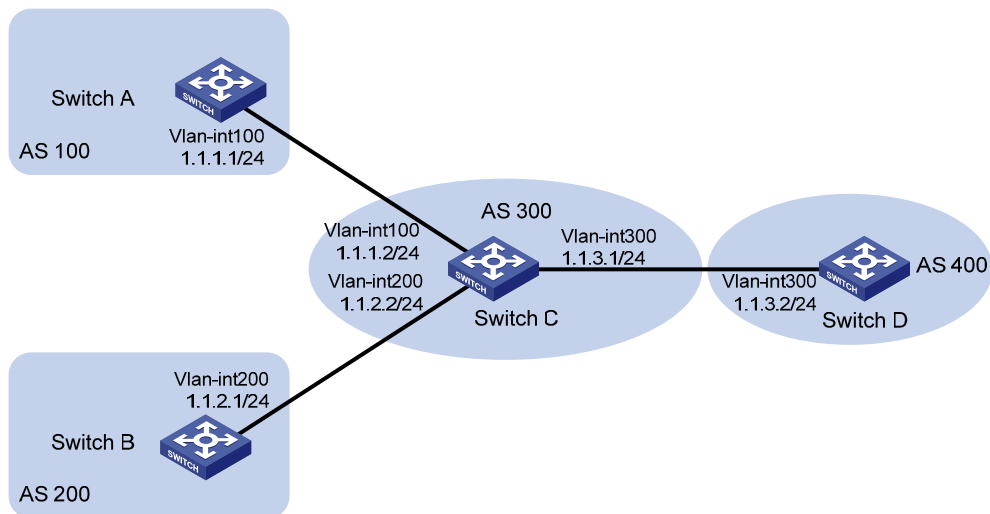
Applying a routing policy to filter received BGP routes

Network requirements

As shown in Figure 121, all the switches run BGP. Switch C establishes eBGP connections with other switches.

Configure a routing policy on Switch D to reject routes from AS 200.

Figure 121 Network diagram



Configuration procedure

1. Configure IP addresses for the interfaces. (Details not shown.)
2. Configure BGP:

```

# Configure Switch A.
<SwitchA> system-view
[SwitchA] bgp 100
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] peer 1.1.1.2 as-number 300

```

Configure Switch B.

```
<SwitchB> system-view
[SwitchB] bgp 200
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] peer 1.1.2.2 as-number 300
```

Configure Switch C.

```
<SwitchC> system-view
[SwitchC] bgp 300
[SwitchC-bgp] router-id 3.3.3.3
[SwitchC-bgp] peer 1.1.1.1 as-number 100
[SwitchC-bgp] peer 1.1.2.1 as-number 200
[SwitchC-bgp] peer 1.1.3.2 as-number 400
```

Configure Switch D.

```
<SwitchD> system-view
[SwitchD] bgp 400
[SwitchD-bgp] router-id 4.4.4.4
[SwitchD-bgp] peer 1.1.3.1 as-number 300
[SwitchD-bgp] quit
```

On Switch A, inject routes 4.4.4.4/24, 5.5.5.5/24, and 6.6.6.6/24 to BGP.

```
[SwitchA-bgp] network 4.4.4.4 24
[SwitchA-bgp] network 5.5.5.5 24
[SwitchA-bgp] network 6.6.6.6 24
```

On Switch B, inject routes 7.7.7.7/24, 8.8.8.8/24, and 9.9.9.9/24 to BGP.

```
[SwitchB-bgp] network 7.7.7.7 24
[SwitchB-bgp] network 8.8.8.8 24
[SwitchB-bgp] network 9.9.9.9 24
```

Display the BGP routing table information of Switch D.

```
[SwitchD-bgp] display bgp routing-table
```

Total Number of Routes: 6

BGP Local router ID is 4.4.4.4

Status codes: * - valid, > - best, d - damped,

h - history, i - internal, s - suppressed, S - Stale

Origin : i - IGP, e - EGP, ? - incomplete

	Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*>	4.4.4.0/24	1.1.3.1			0	300 100i
*>	5.5.5.0/24	1.1.3.1			0	300 100i
*>	6.6.6.0/24	1.1.3.1			0	300 100i
*>	7.7.7.0/24	1.1.3.1			0	300 200i
*>	8.8.8.0/24	1.1.3.1			0	300 200i
*>	9.9.9.0/24	1.1.3.1			0	300 200i

The output shows that Switch D has learned routes 4.4.4.0/24, 5.5.5.0/24, and 6.6.6.0/24 from AS 100 and 7.7.7.0/24, 8.8.8.0/24, and 9.9.9.0/24 from AS 200.

3. Configure Switch D to reject routes from AS 200:

```
# Configure AS_PATH list 1.
```



```

[SwitchD] ip as-path 1 permit .*200.*
# Create routing policy rt1 with node 1, and specify the match mode as deny to deny routes from AS 200.
[SwitchD] route-policy rt1 deny node 1
[SwitchD-route-policy] if-match as-path 1
[SwitchD-route-policy] quit
# Create routing policy rt1 with node 10, and specify the match mode as permit to permit routes from other ASs.
[SwitchD] route-policy rt1 permit node 10
[SwitchD-route-policy] quit
# On Switch D, specify routing policy rt1 to filter routes received from peer 1.1.3.1.
[SwitchD] bgp 400
[SwitchD-bgp] peer 1.1.3.1 route-policy rt1 import
# Display the BGP routing table information of Switch D.
[SwitchD-bgp] display bgp routing-table

```

Total Number of Routes: 3

BGP Local router ID is 4.4.4.4

Status codes: * - valid, > - best, d - damped,

h - history, i - internal, s - suppressed, S - Stale

Origin : i - IGP, e - EGP, ? - incomplete

Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
*> 4.4.4.0/24	1.1.3.1			0	300 100i
*> 5.5.5.0/24	1.1.3.1			0	300 100i
*> 6.6.6.0/24	1.1.3.1			0	300 100i

The output shows that Switch D has learned only routes 4.4.4.0/24, 5.5.5.0/24, and 6.6.6.0/24 from AS 100.

Troubleshooting routing policy configuration

IPv4 routing information filtering failure

Symptom

The routing protocol is running properly, but filtering routing information failed.

Analysis

At least one item of the IP prefix list must be configured as permit mode, and at least one node in the routing policy must be configured as permit mode.

Solution

1. Use the **display ip ip-prefix** command to display IP prefix list information.
2. Use the **display route-policy** command to display routing policy information.

IPv6 routing information filtering failure

Symptom

The routing protocol is running properly, but filtering routing information failed.

Analysis

At least one item of the IPv6 prefix list must be configured as permit mode, and at least one node of the routing policy must be configured as permit mode.

Solution

1. Use the **display ip ipv6-prefix** command to display IP prefix list information.
 2. Use the **display route-policy** command to display routing policy information.
-

Configuring policy-based routing

Hardware compatibility

The HP 5500 SI Switch Series does not support PBR.

Introduction to PBR

Policy-based routing (PBR) is a routing mechanism based on user-defined policies. Different from the traditional destination-based routing mechanism, PBR enables you to use a policy (based on the source address or other criteria) to route packets.

In general, PBR takes precedence over destination-based routing. PBR is applied to the packets matching the specified criteria, and other packets are forwarded through destination-based routing.

This Switch Series supports the following PBR implementation methods:

- Using a PBR policy

The PBR policy allows you to specify the next hop, priority, and default next hop to guide the forwarding of packets that match specific ACLs. Only IPv4 unicast PBR is supported.

You can modify ACL rules to implement flexible service control.

NOTE:

- If the configured next hop in the PBR policy does not exist, the matching packet is forwarded according to the routing table.
 - If a default next hop is configured in the PBR policy, destination-based routing takes precedence over PBR.
-

- Using a QoS policy

The QoS policy uses QoS traffic classification to define matching criteria, and uses the redirection action of traffic behavior to guide packet forwarding. This implementation method supports both IPv4 and IPv6 unicast PBR.

PBR modes

According to the objects to which the PBR applies, PBR involves local PBR and interface PBR.

- Local PBR applies to locally generated packets only, such as the ICMP packets generated by using the **ping** command.
- Interface PBR applies to packets forwarded through the interface only.

In most cases, interface PBR is implemented to meet the forwarding and security requirements.

NOTE:

The term "interface" in this chapter refers to Layer 3 interfaces, including VLAN interfaces and route mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).

Concepts

Policy

A policy is used to route IP packets. A policy can consist of one or multiple nodes.

Node

A node is identified by a node number. The node with the smallest node number has the highest priority.

A policy node consists of **if-match** and **apply** clauses. An **if-match** clause specifies a match criterion on a node, and an **apply** clause specifies an action to be taken on packets.

The action to be taken on matched packets depends on the match mode, which can be **permit** or **deny**.

if-match clause

This Switch Series supports the **if-match acl** clause.

You can specify only one **if-match** clause of each type in a policy node.

apply clause

This Switch Series supports the following **apply** clauses: **apply ip-precedence**, **apply ip-address next-hop**, and **apply ip-address default next-hop**.

NOTE:

The **apply ip-address default next-hop** command takes effect only when no next hop is defined, or the defined next hop is invalid and the destination address does not match any route in the routing table.

Relationship between the match mode and clauses

If a packet...	Then...	
	In permit mode	In deny mode
Matches all the if-match clauses on a policy node	The apply clause is executed, and the packet will not go to the next policy node for a match.	The apply clause is not executed, the packets will not go to the next policy node for a match, and will be forwarded according to the routing table.

If a packet...	Then...	
	In permit mode	In deny mode
Fails to match an if-match clause on the policy node	The apply clause is not executed, and the packet will go to the next policy node for a match.	The apply clause is not executed, and the packet will go to the next policy node for a match.

The nodes of a policy are in an OR relationship. If a packet matches a node, it passes the policy; if the packet does not match any node of the policy, it fails to pass the policy and is forwarded according to the routing table.

QoS mode

The QoS policy uses QoS traffic classification to define matching criteria, and uses the redirection action of traffic behavior to guide packet forwarding in order to implement flexible routing.

PBR takes precedence over destination-based routing. If a packet meets the match criteria, PBR applies; otherwise, destination-based routing applies.

For more information about QoS policies, see *ACL and QoS Configuration Guide*.

Configuring PBR (using a PBR policy)

Defining a policy

Follow these guidelines when you define a policy:

- If an ACL match criterion is defined, packets are matched against the ACL rules, whereas the permit or deny action and the time range of the specified ACL are ignored. If the specified ACL does not exist, no packet is matched.
- You can configure two next hops by using the **apply ip-address next-hop** command twice (first case) or once (second case). After that, executing the **apply ip-address next-hop** command with a new next hop will replace the earlier configured next hop in the first case, or will replace the second next hop specified in the second case. To remove both next hops, execute the **apply ip-address next-hop** command again by specifying two next hops.
- If a policy has a node with no **if-match** clause configured, all packets can pass the policy node. However, an action is taken according to the match mode, and the packets will not go to the next policy node for a match.
- If a policy has a node with the **permit** match mode but no **apply** clause configured, all packets matching all the **if-match** clauses can pass the policy node. However, no action is taken, the packets will not go to the next policy node for a match, and will be forwarded according to the routing table.
- If a policy has a node with no **if-match** or **apply** clauses configured, all packets can pass the policy node. However, no action is taken; the packets will not go to the next policy node for a match, and will be forwarded according to the routing table.

To define a policy:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Create a policy or policy node and enter PBR policy node view.	policy-based-route <i>policy-name</i> [deny permit] node <i>node-number</i>	N/A
3. Define an ACL match criterion.	if-match acl <i>acl-number</i>	Optional.
4. Set an IP precedence type/value.	apply ip-precedence { <i>type</i> <i>value</i> }	Optional.
5. Set next hops.	apply ip-address next-hop <i>ip-address</i> [direct] [track <i>track-entry-number</i>] [<i>ip-address</i> [direct] [track <i>track-entry-number</i>]]	Optional. You can specify two next hops at a time. <ul style="list-style-type: none"> For local PBR, both the two next hops take effect to implement load sharing. For interface PBR, the first next hop serves as the main next hop and the second one serves as the backup next hop.
6. Set default next hops.	apply ip-address default next-hop <i>ip-address</i> [track <i>track-entry-number</i>] [<i>ip-address</i> [track <i>track-entry-number</i>]]	Optional. You can specify two next hops at a time. <ul style="list-style-type: none"> For local PBR, both the two next hops take effect to implement load sharing. For interface PBR, the first next hop serves as the main next hop and the second one serves as the backup next hop.

Configuring local PBR

Only one policy can be referenced for local PBR.

To configure local PBR:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure local PBR based on a policy.	ip local policy-based-route <i>policy-name</i>	Not configured by default.

NOTE:

If the specified policy does not exist, the local PBR configuration succeeds, but it takes effect only when the policy is created.

Configuring interface PBR

Only one policy can be referenced by an interface for interface PBR.

To configure interface PBR:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure interface PBR based on a policy.	ip policy-based-route <i>policy-name</i>	Not configured by default.

NOTE:

If the specified policy does not exist, the interface PBR configuration succeeds, but it takes effect only when the policy is created.

PBR and track

Associated with a Track object, PBR can sense topology changes faster. You can associate PBR with a track entry when configuring the outgoing interface, default outgoing interface, next hop, and default next hop to dynamically determine link reachability. The PBR configuration takes effect when the status of the associated Track object is **Positive** or **Invalid**.

For more information about track-PBR collaboration, see *High Availability Configuration Guide*.

Configuring PBR (using a QoS policy)

Before you configure PBR, complete the following tasks:

- Configure a QoS policy by configuring the match criteria and a redirection action for PBR.
- Apply the QoS policy by defining the occasion to which the PBR applies.

Configuring a QoS policy

To configure a QoS policy:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a class and enter class view.	traffic classifier <i>tcl-name</i> [operator { and or }]	N/A
3. Configure the match criteria.	if-match <i>match-criteria</i>	N/A
4. Exit class view.	quit	N/A
5. Create a behavior and enter behavior view.	traffic behavior <i>behavior-name</i>	N/A
6. Configure a traffic redirection action.	redirect next-hop { <i>ipv4-add1</i> [<i>ipv4-add2</i>] <i>ipv6-add1</i> [<i>interface-type interface-number</i>] [<i>ipv6-add2</i> [<i>interface-type interface-number</i>]] } [fail-action { discard forward }]	Optional. The default option for the fail-action is forward .

Step	Command	Remarks
7. Exit behavior view.	quit	N/A
8. Create a policy and enter policy view.	qos policy <i>policy-name</i>	N/A
9. Associate the class with the traffic behavior in the QoS policy.	classifier <i>tcl-name</i> behavior <i>behavior-name</i>	N/A

Applying the QoS policy

When configuring PBR, you can apply a QoS policy to the following occasions:

- **Applied globally**—Affects the traffic sent or received on all ports.
- **Applied to an interface**—Affects the traffic sent or received on the interface.
- **Applied to a VLAN**—Affects the traffic sent or received on all ports in the VLAN.

NOTE:

A QoS policy used for PBR applies only to traffic received on all ports, on an interface, or on all ports in a VLAN.

To apply the QoS policy globally:

Step	Command
1. Enter system view.	system-view
2. Apply the QoS policy globally.	qos apply policy <i>policy-name</i> global inbound

To apply the QoS policy to an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> • Enter interface view: interface <i>interface-type</i> <i>interface-number</i> • Enter port group view: port-group manual <i>port-group-name</i> 	Use either approach. Settings in interface view take effect on the current interface; settings in port group view take effect on all ports in the port group.
3. Apply the policy to the interface or port group.	qos apply policy <i>policy-name</i> inbound	N/A

To apply the QoS policy to a VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Apply the QoS policy to VLANs.	qos vlan-policy <i>policy-name</i> vlan <i>vlan-id-list</i> inbound	N/A

NOTE:

QoS policies cannot be applied to dynamic VLANs, for example, VLANs created by GVRP.

Displaying and maintaining PBR configuration

PBR configuration (using a PBR policy)

Task	Command	Remarks
Display the PBR routing information.	display ip policy-based-route [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the specified PBR routing information.	display ip policy-based-route setup { interface <i>interface-type interface-number</i> local <i>policy-name</i> } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display PBR statistics.	display ip policy-based-route statistics { interface <i>interface-type interface-number</i> local } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the PBR policy information.	display policy-based-route [<i>policy-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear PBR statistics.	reset policy-based-route statistics [<i>policy-name</i>]	Available in user view

If a policy has a node with no **if-match** or **apply** clause configured, all packets can pass the policy. However, no action is taken and the packets will not go to the next policy node for a match. The statistics of PBR will be changed.

If a policy node has **if-match** clauses, but no **apply** clauses configured, packets will match against these **if-match** clauses. However, no **apply** clauses are applicable to the permitted packets, and the packets will not go to the next policy node for a match. The statistics of PBR will be changed.

If a policy node has no **if-match** clause, but **apply** clauses configured, all packets can pass the policy, and then are forwarded according to the **apply** clauses if the **permit** keyword is specified for the node, or are denied if the **deny** keyword is specified. The packets will not go to the next policy node for a match. The statistics of PBR will be changed.

If the match mode of a policy node is **deny**, no **apply** clause will be executed for the packets satisfying all the **if-match** clauses, and the packets will not go to the next policy node for a match. They will be forwarded according to the routing table instead. Neither debugging information nor statistics for the **deny** match mode can be displayed.

PBR configuration (using a QoS policy)

Task	Command	Remarks
Display user-defined QoS policy configuration information.	display qos policy user-defined [<i>policy-name</i>] [classifier <i>tcl-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display QoS policy configuration on the specified interface or on all interfaces.	display qos policy interface [<i>interface-type interface-number</i>] [inbound outbound] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display VLAN QoS policy information.	display qos vlan-policy { name <i>policy-name</i> vlan <i>vlan-id</i> } [slot <i>slot-number</i>] [inbound outbound] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display information about global QoS policies.	display qos policy global [slot <i>slot-number</i>] [inbound outbound] [[{ begin exclude include } <i>regular-expression</i>]]	Available in user view

PBR configuration examples

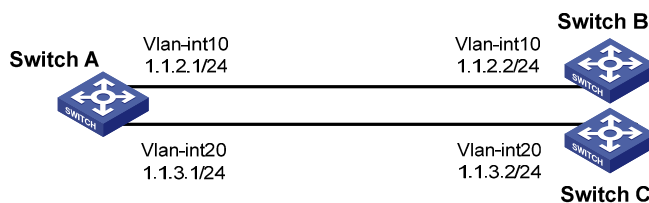
Configuring local PBR based on packet type

Network requirements

As shown in [Figure 122](#), configure PBR on Switch A, so that all TCP packets are forwarded to next hop 1.1.2.2 and other packets are forwarded according to the routing table.

Switch A is directly connected to Switch B and Switch C. Switch B and Switch C are unreachable to each other.

Figure 122 Network diagram



Configuration procedure

1. Configure Switch A:

Define ACL 3101 to match TCP packets.

```

<SwitchA> system-view
[SwitchA] acl number 3101
[SwitchA-acl-adv-3101] rule permit tcp
[SwitchA-acl-adv-3101] quit
  
```

Configure Node 5 of policy **aaa** to forward TCP packets to next hop 1.1.2.2.

```

[SwitchA] policy-based-route aaa permit node 5
[SwitchA-pbr-aaa-5] if-match acl 3101
[SwitchA-pbr-aaa-5] apply ip-address next-hop 1.1.2.2
[SwitchA-pbr-aaa-5] quit
  
```

Apply policy **aaa** to Switch A.

```

[SwitchA] ip local policy-based-route aaa
  
```

Configure the IP addresses of VLAN-interface 10 and VLAN-interface 20.

```

[SwitchA] interface vlan-interface 10
  
```

```
[SwitchA-Vlan-interface10] ip address 1.1.2.1 255.255.255.0
[SwitchA-Vlan-interface10] quit
[SwitchA] interface vlan-interface 20
[SwitchA-Vlan-interface20] ip address 1.1.3.1 255.255.255.0
```

2. Configure Switch B:

Configure the IP address of VLAN-interface 10.

```
<SwitchB> system-view
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] ip address 1.1.2.2 255.255.255.0
[SwitchB-Vlan-interface10] quit
```

3. Configure Switch C:

Configure the IP address of VLAN-interface 20.

```
<SwitchC> system-view
[SwitchC] interface vlan-interface 20
[SwitchC-Vlan-interface20] ip address 1.1.3.2 255.255.255.0
[SwitchC-Vlan-interface20] quit
```

4. Verify the configuration:

Telnet to Switch B (1.1.2.2/24) from Switch A. The operation succeeds.

Telnet to Switch C (1.1.3.2/24) from Switch A. The operation fails.

```
<SwitchA> telnet 1.1.3.2
```

```
Trying 1.1.3.2 ...
```

```
Press CTRL+K to abort
```

```
Can't connect to the remote host!
```

Ping Switch C (1.1.3.2/24) from Switch A. The operation succeeds.

```
<SwitchA> ping 1.1.3.2
```

```
  PING 1.1.3.2: 56 data bytes, press CTRL_C to break
    Reply from 1.1.3.2: bytes=56 Sequence=1 ttl=255 time=2 ms
    Reply from 1.1.3.2: bytes=56 Sequence=2 ttl=255 time=1 ms
    Reply from 1.1.3.2: bytes=56 Sequence=3 ttl=255 time=1 ms
    Reply from 1.1.3.2: bytes=56 Sequence=4 ttl=255 time=1 ms
    Reply from 1.1.3.2: bytes=56 Sequence=5 ttl=255 time=1 ms
--- 1.1.3.2 ping statistics ---
  5 packet(s) transmitted
  5 packet(s) received
  0.00% packet loss
  round-trip min/avg/max = 1/1/2 ms
```

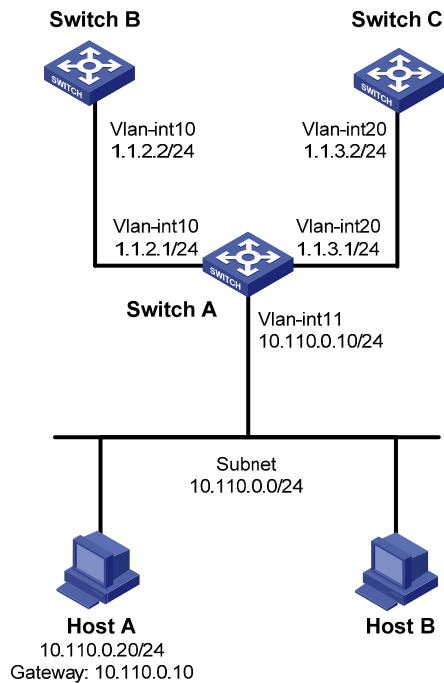
Telnet uses TCP, and ping uses ICMP. The preceding results show that all TCP packets of Switch A are forwarded to next hop 1.1.2.2, and other packets are forwarded via VLAN-interface 20. The PBR configuration is effective.

Configuring interface PBR based on packet type

Network requirements

As shown in [Figure 123](#), configure PBR on Switch A, so that TCP packets arriving on VLAN-interface 11 are forwarded to next hop 1.1.2.2 and other packets are forwarded according to the routing table.

Figure 123 Network diagram



Configuration procedure

In this example, static routes are configured to ensure the reachability among devices.

1. Configure Switch A:

Define ACL 3101 to match TCP packets.

```
<SwitchA> system-view
[SwitchA] acl number 3101
[SwitchA-acl-adv-3101] rule permit tcp
[SwitchA-acl-adv-3101] quit
```

Configure Node 5 of policy **aaa** to forward TCP packets to next hop 1.1.2.2.

```
[SwitchA] policy-based-route aaa permit node 5
[SwitchA-pbr-aaa-5] if-match acl 3101
[SwitchA-pbr-aaa-5] apply ip-address next-hop 1.1.2.2
[SwitchA-pbr-aaa-5] quit
```

Apply the policy **aaa** to VLAN-interface 11.

```
[SwitchA] interface vlan-interface 11
[SwitchA-Vlan-interface11] ip address 10.110.0.10 255.255.255.0
[SwitchA-Vlan-interface11] ip policy-based-route aaa
[SwitchA-Vlan-interface11] quit
```

Configure the IP addresses of VLAN-interface 10 and VLAN-interface 20.

```
[SwitchA] interface vlan-interface 10
[SwitchA-Vlan-interface10] ip address 1.1.2.1 255.255.255.0
[SwitchA-Vlan-interface10] quit
[SwitchA] interface vlan-interface 20
[SwitchA-Vlan-interface20] ip address 1.1.3.1 255.255.255.0
```

2. Configure Switch B:

Configure a static route to subnet 10.110.0.0/24.

```

<SwitchB> system-view
[SwitchB] ip route-static 10.110.0.0 24 1.1.2.1
# Configure the IP address of VLAN-interface 10.
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] ip address 1.1.2.2 255.255.255.0
[SwitchB-Vlan-interface10] quit

```

3. Configure Switch C:

```

# Configure a static route to subnet 10.110.0.0/24.
<SwitchC> system-view
[SwitchC] ip route-static 10.110.0.0 24 1.1.3.1
# Configure the IP address of VLAN-interface 20.
[SwitchC] interface vlan-interface 20
[SwitchC-Vlan-interface20] ip address 1.1.3.2 255.255.255.0
[SwitchC-Vlan-interface20] quit

```

Verifying the configuration

Configure the IP address of Host A as 10.110.0.20/24, and the gateway as 10.110.0.10.

On Host A, Telnet to Switch B (1.1.2.2) that is directly connected to Switch A. The operation succeeds.

On Host A, Telnet to Switch C (1.1.3.2) that is directly connected to Switch A. The operation fails.

Ping Switch C from Host A. The operation succeeds.

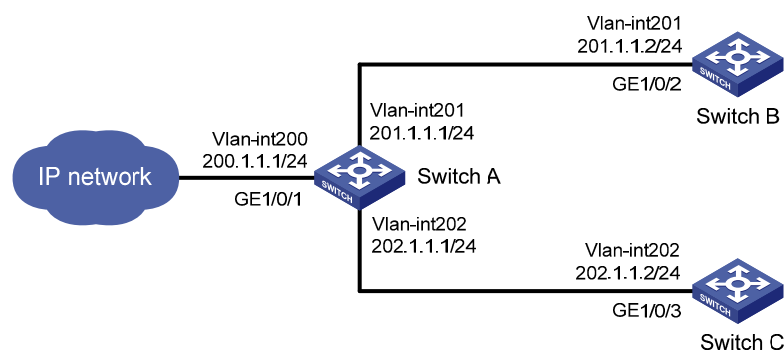
Telnet uses TCP, and ping uses ICMP. The preceding results show that all TCP packets arriving on VLAN-interface 11 of Switch A are forwarded to next hop 1.1.2.2, and other packets are forwarded via VLAN-interface 20. The PBR configuration is effective.

IPv4 PBR configuration example (using a QoS policy)

Network requirements

As shown in [Figure 124](#), redirect all packets received on GigabitEthernet 1/0/1 of Switch A to the next hop 202.1.1.2.

Figure 124 Network diagram



Configuration procedure

```

# Configure ACL 2000.
<SwitchA> system-view
[SwitchA] acl number 2000

```

```
[SwitchA-acl-basic-2000] rule 0 permit source any
[SwitchA-acl-basic-2000] quit

# Define a match criterion for class a to match ACL 2000.
[SwitchA] traffic classifier a
[SwitchA-classifier-a] if-match acl 2000
[SwitchA-classifier-a] quit

# Configure the action of redirecting traffic to the next hop 202.1.1.2 for behavior a.
[SwitchA] traffic behavior a
[SwitchA-behavior-a] redirect next-hop 202.1.1.2
[SwitchA-behavior-a] quit

# Associate class a with behavior a in QoS policy a.
[SwitchA] qos policy a
[SwitchA-qospolicy-a] classifier a behavior a
[SwitchA-qospolicy-a] quit

# Apply QoS policy a to the incoming traffic of GigabitEthernet 1/0/1.
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] qos apply policy a inbound
```

Verifying the configuration

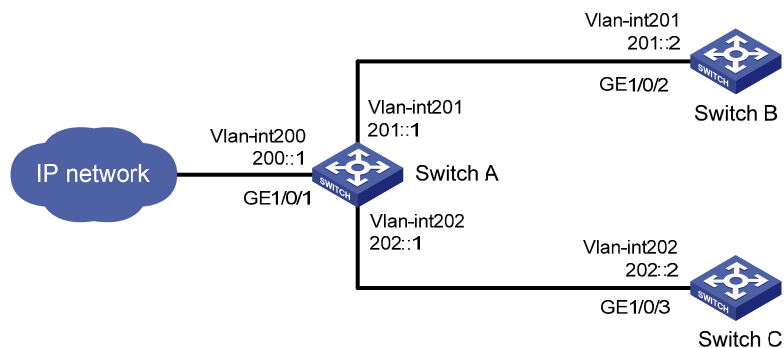
After completing the configuration, verify that when Switch A receives packets with destination IP address 201.1.1.2, it forwards the packets to Switch C instead of Switch B.

IPv6 PBR configuration example (using a QoS policy)

Network requirements

As shown in [Figure 125](#), redirect all packets received on GigabitEthernet 1/0/1 of Switch A to the next hop 202::2.

Figure 125 Network diagram



Configuration procedure

```
# Configure IPv6 ACL 2000.
<SwitchA> system-view
[SwitchA] acl ipv6 number 2000
[SwitchA-acl6-basic-2000] rule 0 permit source any
[SwitchA-acl6-basic-2000] quit

# Define a match criterion for class a to match IPv6 ACL 2000.
```

```
[SwitchA] traffic classifier a
[SwitchA-classifier-a] if-match acl ipv6 2000
[SwitchA-classifier-a] quit

# Configure the action of redirecting traffic to the next hop 202::2 for behavior a.
[SwitchA] traffic behavior a
[SwitchA-behavior-a] redirect next-hop 202::2
[SwitchA-behavior-a] quit

# Associate class a with behavior a in QoS policy a.
[SwitchA] qos policy a
[SwitchA-qospolicy-a] classifier a behavior a
[SwitchA-qospolicy-a] quit

# Apply QoS policy a to the incoming traffic of GigabitEthernet 1/0/1.
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] qos apply policy a inbound
```

Verifying th configuration

After completing the configuration, verify that when Switch A receives packets with destination IP address 201::2, it forwards the packets to Switch C instead of Switch B.

Configuring MCE

The term "router" in this document refers to both routers and Layer 3 switches.

The term "Layer 3 interface" in this chapter refers to route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).

This chapter covers MCE related configuration. For information about routing protocols, see *Layer 3—IP Services Configuration Guide*. The MCE function is available only on the HP 5500 EI switch series.

MCE overview

MPLS L3VPN overview

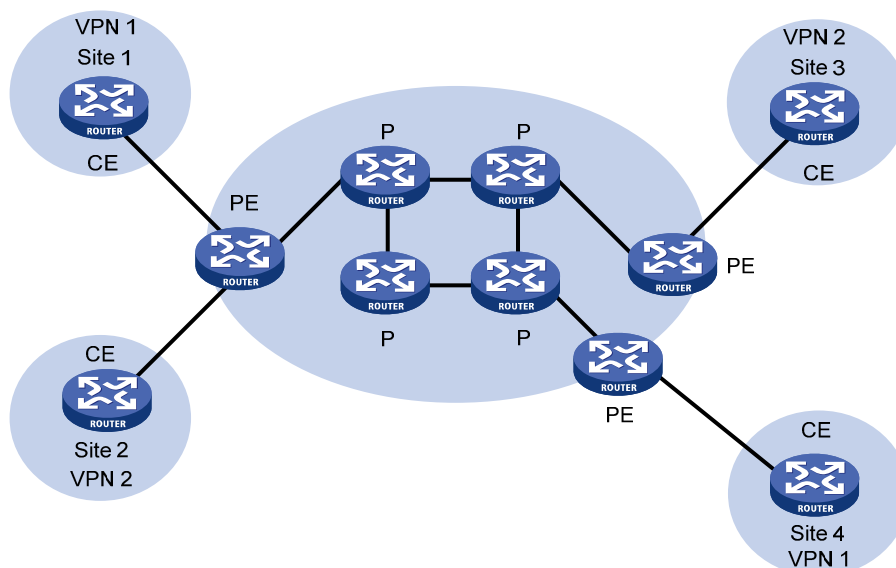
MPLS L3VPN is a type of PE-based L3VPN technology for service provider VPN solutions. It uses BGP to advertise VPN routes and uses MPLS to forward VPN packets on service provider backbones.

MPLS L3VPN provides flexible networking modes, excellent scalability, and convenient support for MPLS QoS and MPLS TE.

The MPLS L3VPN model consists of the following types of devices:

- **Customer edge (CE) device**—A CE resides on a customer network and has one or more interfaces directly connected with service provider networks. It can be a router, a switch, or a host. It can neither "sense" the existence of any VPN nor does it need to support MPLS.
- **Provider edge (PE) device**—A PE resides on a service provider network and connects one or more CEs to the network. On an MPLS network, all VPN processing occurs on the PEs.
- **Provider (P) device**—A P device is a core device on a service provider network. It is not directly connected with any CE. It only needs to be equipped with basic MPLS forwarding capability.

Figure 126 Network diagram for MPLS L3VPN model



CEs and PEs mark the boundary between the service providers and the customers.

After a CE establishes adjacency with a directly connected PE, it advertises its VPN routes to the PE and learns remote VPN routes from the PE. A CE and a PE use BGP/IGP to exchange routing information. You can also configure static routes between them.

After a PE learns the VPN routing information of a CE, it uses BGP to exchange VPN routing information with other PEs. A PE maintains routing information about only VPNs that are directly connected, rather than all VPN routing information on the provider network.

A P router maintains only routes to PEs and does not deal with VPN routing information.

When VPN traffic travels over the MPLS backbone, the ingress PE functions as the ingress Label Switching Router (LSR), the egress PE functions as the egress LSR, and P routers function as the transit LSRs.

MPLS L3VPN concepts

Site

Sites are often mentioned in the VPN. A site has the following features:

- A site is a group of IP systems with IP connectivity that does not rely on any service provider network to implement.
- The classification of a site depends on the topology relationship of the devices, rather than the geographical positions, though the devices at a site are, in most cases, adjacent to each other geographically.
- The devices at a site can belong to multiple VPNs.
- A site is connected to a provider network through one or more CEs. A site can contain many CEs, but a CE can belong to only one site.

Sites connected to the same provider network can be classified into different sets by policies. Only the sites in the same set can access each other through the provider network. Such a set is called a VPN.

Address space overlapping

Each VPN independently manages the addresses it uses. The assembly of such addresses for a VPN is called an address space.

The address spaces of VPNs may overlap. For example, if both VPN 1 and VPN 2 use the addresses on network segment 10.110.10.0/24, address space overlapping occurs.

VPN instance

In MPLS VPN, routes of different VPNs are identified by VPN instance.

A PE creates and maintains a separate VPN instance for each VPN at a directly connected site. Each VPN instance contains the VPN membership and routing rules of the corresponding site. If a user at a site belongs to multiple VPNs at the same time, the VPN instance of the site contains information about all of the VPNs.

For independence and security of VPN data, each VPN instance on a PE maintains a relatively independent routing table and a separate label forwarding information base (LFIB). VPN instance information contains the following items: the LFIB, IP routing table, interfaces bound to the VPN instance, and administration information of the VPN instance. The administration information of the VPN instance includes the route distinguisher (RD), route filtering policy, and member interface list.

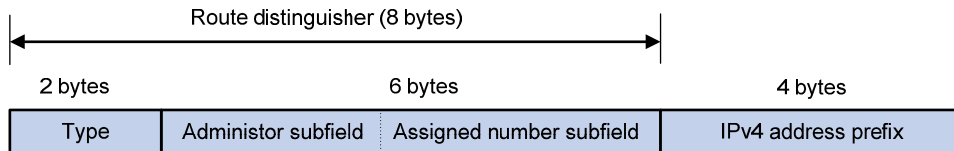
VPN-IPv4 address

Traditional BGP cannot process overlapping VPN routes. If, for example, both VPN 1 and VPN 2 use addresses on the segment 10.110.10.0/24 and each advertise a route to the segment, BGP selects only one of them, which results in the loss of the other route.

PEs use MP-BGP to advertise VPN routes and use VPN-IPv4 address family to solve the problem with traditional BGP.

A VPN-IPv4 address consists of 12 bytes. The first eight bytes represent the RD, followed by a four-byte IPv4 address prefix.

Figure 127 VPN-IPv4 address structure



When a PE receives an ordinary IPv4 route from a CE, it must advertise the VPN route to the peer PE. The uniqueness of a VPN route is implemented by adding an RD to the route.

A service provider can independently assign RDs if the assigned RDs are unique. A PE can advertise different routes to VPNs even if the VPNs are from different service providers and are using the same IPv4 address space.

Configure a distinct RD for each VPN instance on a PE, so that routes to the same CE use the same RD. The VPN-IPv4 address with an RD of 0 is a globally unique IPv4 address.

By prefixing a distinct RD to a specific IPv4 address prefix, you get a globally unique VPN IPv4 address prefix.

An RD can be related to an autonomous system (AS) number, in which case it is the combination of the AS number and a discretionary number; or it can be related to an IP address, in which case it is the combination of the IP address and a discretionary number.

An RD can be in one of the following formats distinguished by the Type field:

- When the value of the Type field is 0, the Administrator subfield occupies two bytes, the Assigned number subfield occupies four bytes, and the RD format is *16-bit AS number:32-bit user-defined number*. For example, 100:1.
- When the value of the Type field is 1, the Administrator subfield occupies four bytes, the Assigned number subfield occupies two bytes, and the RD format is *32-bit IPv4 address:16-bit user-defined number*. For example, 172.1.1.1:1.
- When the value of the Type field is 2, the Administrator subfield occupies four bytes, the Assigned number subfield occupies two bytes, and the RD format is *32-bit AS number:16-bit user-defined number*, where the minimum value of the AS number is 65536. For example, 65536:1.

To guarantee global uniqueness for an RD, do not set the Administrator subfield to any private AS number or private IP address.

Route target attributes

MPLS L3VPN uses the BGP extended community attributes called "route target" attributes to control the advertisement of VPN routing information.

A VPN instance on a PE supports the following types of route target attributes:

- **Export target attribute:** A local PE sets this type of route target attribute for VPN-IPv4 routes learned from directly connected sites before advertising them to other PEs.
- **Import target attribute:** A PE checks the export target attribute of VPN-IPv4 routes advertised by other PEs. If the export target attribute matches the import target attribute of the VPN instance, the PE adds the routes to the VPN routing table.

In other words, route target attributes define which sites can receive VPN-IPv4 routes, and from which sites that a PE can receive routes.

Similar to RDs, route target attributes can be of the following formats:

- *16-bit AS number:32-bit user-defined number.* For example, 100:1.
- *32-bit IPv4 address: 16-bit user-defined number.* For example, 172.1.1.1:1.
- *32-bit AS number: 16-bit user-defined number,* where the minimum value of the AS number is 65536. For example, 65536:1.

Multi-VPN-instance CE

Using tunnels, MPLS L3VPN implements private network data transmission over the public network. However, the traditional MPLS L3VPN architecture requires each VPN instance exclusively use a CE to connect with a PE, as shown in [Figure 126](#).

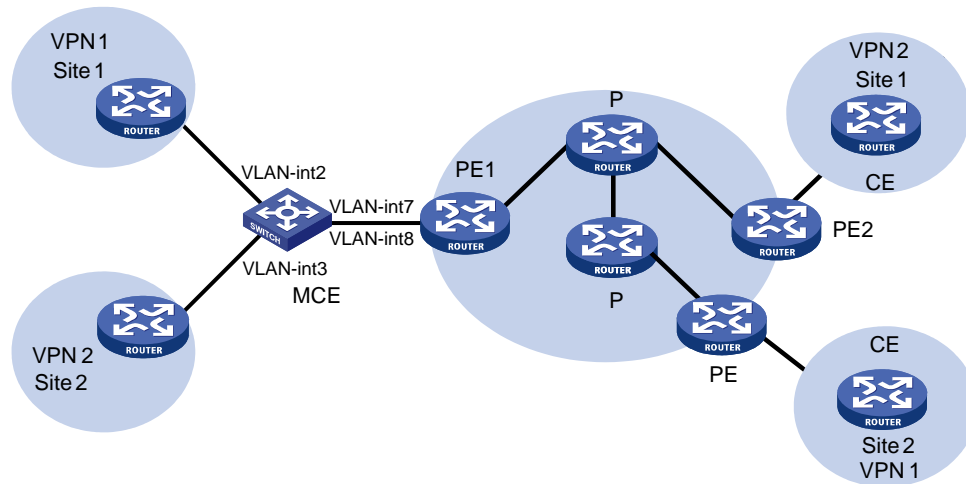
For better services and higher security, a private network is usually divided into multiple VPNs to isolate services. To meet these requirements, you can configure a CE for each VPN, which increases users' device expenses and maintenance costs. Or, you can configure multiple VPNs to use the same CE and the same routing table, which sacrifices data security.

Using the Multi-VPN-Instance CE (MCE) function of the switch, you can remove the contradiction of low cost and high security in multi-VPN networks. With MCE configured, a CE can bind each VPN in a network with a VLAN interface on the CE, and create and maintain a separate routing table (multi-VRF) for each VPN. This separates the forwarding paths of packets of different VPNs and, in conjunction with the PE, can correctly advertise the routes of each VPN to the peer PE, ensuring the normal transmission of VPN packets over the public network.

How MCE works

[Figure 128](#) shows how an MCE maintains the routing entries of multiple VPNs and how an MCE exchanges VPN routes with PEs.

Figure 128 Network diagram for the MCE function



On the left-side network, there are two VPN sites, both of which are connected to the MPLS backbone through the MCE device. VPN 1 and VPN 2 on the left-side network must establish a tunnel with VPN 1 and VPN 2 on the right-side network, respectively.

With MCE enabled, routing tables can be created for VPN 1 and VPN 2 individually, VLAN-interface 2 can be bound to VPN 1, and VLAN-interface 3 can be bound to VPN 2. When receiving a piece of routing information, MCE determines the source of the routing information according to the number of the interface receiving the information. It then maintains the corresponding routing table accordingly.

You must also bind the interfaces to the VPNs on PE 1 in the same way as those on the MCE device. The MCE device is connected to PE 1 through a trunk, which permits packets of VLAN 2 and VLAN 3 with VLAN tags carried. In this way, PE 1 can determine the VPN a received packet belongs to according to the VLAN tag of the packet and passes the packet to the corresponding tunnel.

Configuring routing on an MCE

Interface-to-VPN-instance binding enables MCEs and PEs to determine the sources of received packets and then forward the packets according to the routing information concerning the corresponding VPNs. MCE routing configuration includes:

- MCE-VPN site routing configuration
- MCE-PE routing configuration

Route exchange between an MCE and a VPN site

An MCE can adopt the following routing protocols to exchange VPN routes with a site:

- Static route
- RIP
- OSPF
- IS-IS
- IBGP
- EBGp

This section briefly introduces the cooperation of routing protocols and MCE.

Static routes

An MCE can communicate with a site through static routes. As static routes configured for traditional CEs take effect globally, address overlapping between multiple VPNs remains a problem until the emergence of MCE. MCE allows static-route-to-VPN-instance binding, which isolates the static routes of different VPNs.

RIP

The switch can bind RIP processes to VPN instances. With these bindings on the MCE, private network routes of different VPNs can be exchanged between MCE and sites through different RIP processes, isolating and securing VPN routes.

OSPF

The switch can bind OSPF processes to VPN instances and isolate the routes of different VPNs.

For an OSPF process bound to a VPN instance, the router ID of the public network configured in system view is invalid. You must specify the router ID when creating an OSPF process.

An OSPF process can be bound to only one VPN instance. However, a VPN instance can use multiple OSPF processes for private network route transmission. To make sure routes can be advertised properly, configure the same domain ID for all OSPF processes bound to the same VPN instance.

Routes redistributed from OSPF to BGP on the MCE have their OSPF attributes removed. To enable BGP to distinguish routes redistributed from different OSPF domains, you must enable the redistributed routes to carry the OSPF domain ID by configuring the **domain-id** command in OSPF view. The domain ID is added to BGP VPN routes as an extended community attribute.

In cases where a VPN has multiple MCE devices attached to it and when an MCE device advertises the routes learned from BGP within the VPN, the routes may be learned by other MCE devices, generating route loops. To prevent route loops, configure route tags for different VPN instances on each MCE. HP recommends that you assign the same route tag to the same VPN on all MCEs.

IS-IS

Similar to those in OSPF, IS-IS processes can be bound to VPN instances for private network routes to be exchanged between MCE and sites. An IS-IS process can be bound to only one VPN instance.

IBGP

To use IBGP to exchange private routes between an MCE and a site, configure IBGP peers for VPN instances on the MCE and redistribute IGP routing information from corresponding VPNs. If the MCE is connected with multiple sites in the same VPN, you can configure the MCE as a route reflector (RR) and configure the egress routers of the sites as clients, making the MCE reflect routing information between the sites. This eliminates the necessity for BGP connections between sites, reducing the number of BGP connections and simplifying network configuration.

EBGP

To use EBGP for exchanging routing information between an MCE and VPN sites, you must configure a BGP peer for each VPN instance on the MCE, and redistribute the IGP routes of each VPN instance on the VPN sites. You also can configure filtering policies to filter the received routes and the routes to be advertised.

Route exchange between an MCE and a PE

Routing information entries are bound to specific VPN instances on an MCE device, and packets of each VPN instance are forwarded between MCE and PE according to interface. As a result, VPN routing

information can be transmitted by performing relatively simple configurations between MCE and PE, such as importing the VPN routing entries on MCE devices to the routing table of the routing protocol running between MCE and PEs.

The following routing protocols can be used between MCE and PE devices for routing formation exchange:

- Static route
- RIP
- OSPF
- IS-IS
- IBGP
- EBGp

Configuring an MCE

Configuring VPN instances

Configuring VPN instances is required in all MCE networking schemes.

By configuring VPN instances on a PE, you isolate not only VPN routes from public network routes, but also routes of a VPN from those of another VPN. This feature allows VPN instances to be used in networking scenarios besides MCE.

Creating a VPN Instance

A VPN instance is associated with a site. It is a collection of the VPN membership and routing rules of its associated site. A VPN instance does not necessarily correspond to one VPN.

A VPN instance takes effect only after you configure an RD for it. Before configuring an RD for a VPN instance, you can configure no other parameters for the instance but a description.

You can configure a description for a VPN instance to record its related information, such as its relationship with a certain VPN.

To create and configure a VPN instance:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a VPN instance and enter VPN instance view.	ip vpn-instance <i>vpn-instance-name</i>	N/A
3. Configure an RD for the VPN instance.	route-distinguisher <i>route-distinguisher</i>	For easy management, set the same RD for the same VPN instance on the MCE and PE.
4. Configure a description for the VPN instance.	description <i>text</i>	Optional

Associating a VPN instance with an interface

In an MPLS L3VPN application, you must associate VPN instances with the interfaces connecting the PEs.

In a tunneling application, you must associate VPN instances with the tunnel interfaces connecting the peer MCE devices or CE devices.

You can add a management Ethernet interface on the switch to a VPN so that the IP address of the interface only participates in the route calculation of the specified VPN.

After creating and configuring a VPN instance, you associate the VPN instance with the interface for connecting different VPN sites.

To associate a VPN instance with an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Associate the current interface with the VPN instance.	ip binding vpn-instance <i>vpn-instance-name</i>	No VPN instance is associated with an interface by default.

NOTE:

The **ip binding vpn-instance** command clears the IP address of the interface on which it is configured. Be sure to reconfigure an IP address for the interface after configuring the command.

Configuring route-related attributes of a VPN instance

The control process of VPN route advertisement is as follows:

- When a VPN route learned from a site gets redistributed into BGP, BGP associates it with a route target extended community attribute list, which is usually the export target attribute of the VPN instance associated with the site.
- The VPN instance determines which routes it can accept and redistribute according to the **import-extcommunity** in the route target.
- The VPN instance determines how to change the route targets attributes for routes to be advertised according to the **export-extcommunity** in the route target.

! **IMPORTANT:**

- Only when BGP runs between the MCE and PE can the route target attribute be advertised to the PE along with the routing information. In other cases, configuring this attribute makes no sense.
- Before associating a routing policy with a VPN instance, you must first create the routing policy. Otherwise, the default routing policy is used.

To configure route related attributes of a VPN instance:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VPN instance view.	ip vpn-instance <i>vpn-instance-name</i>	N/A
3. Enter IPv4 VPN view.	ipv4-family	Optional.
4. Associate the current VPN instance with one or more route targets.	vpn-target <i>vpn-target</i> <1-8> [both export-extcommunity import-extcommunity]	A single vpn-target command can configure up to eight route targets. You can configure up to 64 route targets for a VPN instance.

Step	Command	Remarks
5. Configure the maximum number of routes for the VPN instance.	routing-table limit <i>number</i> { <i>warn-threshold</i> simply-alert }	Optional. Not configured by default. Setting the maximum number of routes for a VPN instance to support is for preventing too many routes from being redistributed into the PE.
6. Apply an import routing policy to the current VPN instance.	import route-policy <i>route-policy</i>	Optional. By default, all routes permitted by the import target attribute can be redistributed into the VPN instance.
7. Apply an export routing policy to the current VPN instance.	export route-policy <i>route-policy</i>	Optional. By default, all VPN instance routes permitted by the export target attribute can be redistributed.

NOTE:

- Only when BGP runs between the MCE and PE can the route target attribute be advertised to the PE together with the routing information. In other cases, configuring this attribute makes no sense.
- You can configure route related attributes for IPv4 VPNs in both VPN instance view and IPv4 VPN view. Those configured in IPv4 VPN view take precedence.

Configuring routing on an MCE

MCE implements service isolation through route isolation. MCE routing configuration includes:

- MCE-VPN site routing configuration
- MCE-PE routing configuration

On the PE in an MCE network environment, disable routing loop detection to avoid route loss during route calculation and disable route redistribution between routing protocols to save system resources.

Configuration prerequisites

Before you configure routing on an MCE, complete the following tasks:

- On the MCE, configure VPN instances, and bind the VPN instances with the interfaces connected to the VPN sites and those connected to the PE.
- Configure the link layer and network layer protocols on related interfaces to ensure IP connectivity.

Configuring routing between MCE and VPN site

Configuring static routing between MCE and VPN site

An MCE can reach a VPN site through a static route. Static routing on a traditional CE is globally effective and thus does not support address overlapping among VPNs. An MCE supports binding a static route with a VPN instance, so that the static routes of different VPN instances can be isolated from each other.

To configure static routing between MCE and VPN site:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure a static route for a VPN instance.	<ul style="list-style-type: none"> ip route-static <i>dest-address</i> { <i>mask</i> <i>mask-length</i> } { <i>gateway-address</i> <i>interface-type interface-number</i> [<i>gateway-address</i>] vpn-instance <i>d-vpn-instance-name gateway-address</i> } [preference <i>preference-value</i>] [tag <i>tag-value</i>] [description <i>description-text</i>] ip route-static vpn-instance <i>s-vpn-instance-name</i> <1-6> <i>dest-address</i> { <i>mask</i> <i>mask-length</i> } { <i>gateway-address</i> [public] <i>interface-type interface-number</i> [<i>gateway-address</i>] vpn-instance <i>d-vpn-instance-name gateway-address</i> } [preference <i>preference-value</i>] [tag <i>tag-value</i>] [description <i>description-text</i>] 	Use either command. Perform this configuration on the MCE. On a VPN site, configure a normal static route.
3. Configure the default precedence for static routes.	ip route-static default-preference <i>default-preference-value</i>	Optional. 60 by default.

Configuring RIP between MCE and VPN site

A RIP process belongs to the public network or a single VPN instance. If you create a RIP process without binding it to a VPN instance, the process belongs to the public network. By configuring RIP process-to-VPN instance bindings on a IPv6 MCE, you allow routes of different VPNs to be exchanged between the MCE and the sites through different RIP processes, ensuring the separation and security of VPN routes.

To configure RIP between MCE and VPN site:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a RIP process for a VPN instance and enter RIP view.	rip [<i>process-id</i>] vpn-instance <i>vpn-instance-name</i>	Perform this configuration on the MCE. On a VPN site, create a normal RIP process.
3. Enable RIP on the interface attached to the specified network.	network <i>network-address</i>	By default, RIP is disabled on an interface.
4. Redistribute remote site routes advertised by the PE.	import-route <i>protocol</i> [<i>process-id</i>] [allow-ibgp] [cost <i>cost</i>] route-policy <i>route-policy-name</i> tag <i>tag</i>] *	By default, no route is redistributed into RIP.
5. Configure the default cost value for the redistributed routes.	default cost <i>value</i>	Optional. 0 by default.

Configuring OSPF between MCE and VPN site

An OSPF process belongs to the public network or a single VPN instance. If you create an OSPF process without binding it to a VPN instance, the process belongs to the public network.

By configuring OSPF process-to-VPN instance bindings on a MCE, you allow routes of different VPNs to be exchanged between the MCE and the sites through different OSPF processes, ensuring the separation and security of VPN routes.

An OSPF process can belong to only one VPN instance, but one VPN instance can use multiple OSPF processes to advertise the VPN routes.

An OSPF process that is bound with a VPN instance does not use the public network router ID configured in system view. Therefore, you must configure a router ID when starting the OSPF process. All OSPF processes for the same VPN must be configured with the same OSPF domain ID to ensure correct route advertisement.

To configure OSPF between MCE and VPN site:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create an OSPF process for a VPN instance and enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	Perform this configuration on the MCE. On a VPN site, create a normal OSPF process.
3. Configure the OSPF domain ID.	domain-id <i>domain-id</i> [secondary]	Optional. 0 by default. Perform this configuration on the MCE. On a VPN site, perform the common OSPF configuration.
4. Redistribute remote site routes advertised by the PE.	import-route <i>protocol</i> [<i>process-id</i> allow-ibgp] [cost <i>cost</i> type <i>type</i> tag <i>tag</i> route-policy <i>route-policy-name</i>] *	By default, no route of any other routing protocol is redistributed into OSPF.
5. Create an OSPF area and enter OSPF area view.	area <i>area-id</i>	By default, no OSPF area is created.
6. Enable OSPF on the interface attached to the specified network in the area.	network <i>ip-address</i> <i>wildcard-mask</i>	By default, an interface neither belongs to any area nor runs OSPF.

Configuring IS-IS between MCE and VPN site

An IS-IS process belongs to the public network or a single VPN instance. If you create an IS-IS process without binding it to a VPN instance, the process belongs to the public network.

By configuring IS-IS process-to-VPN instance bindings on a MCE, you allow routes of different VPNs to be exchanged between the MCE and the sites through different IS-IS processes, ensuring the separation and security of VPN routes.

To configure IS-IS between MCE and VPN site:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create an IS-IS process for a VPN instance and enter IS-IS view.	isis [<i>process-id</i>] vpn-instance <i>vpn-instance-name</i>	Perform this configuration on the MCE. On a VPN site, configure a normal IS-IS process.
3. Configure a network entity title.	network-entity <i>net</i>	Not configured by default.

Step	Command	Remarks
4. Redistribute remote site routes advertised by the PE.	import-route { <i>isis</i> [<i>process-id</i>] ospf [<i>process-id</i>] rip [<i>process-id</i>] bgp [allow-ibgp] direct static } [cost <i>cost</i> cost-type { external internal } [level-1 level-1-2 level-2] route-policy <i>route-policy-name</i> tag <i>tag</i>] *	Optional. By default, IS-IS does not redistribute routes of any other routing protocol. If you do not specify the route level in the command, the command will redistribute routes to the level-2 routing table by default.
5. Return to system view.	quit	N/A
6. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
7. Enable the IS-IS process on the interface.	isis enable [<i>process-id</i>]	Disabled by default.

Configuring EBGW between MCE and VPN site

To use EBGW for exchanging routing information between an MCE and VPN sites, you must configure a BGP peer for each VPN instance on the MCE, and redistribute the IGP routes of each VPN instance on the VPN sites.

If EBGW is used for route exchange, you also can configure filtering policies to filter the received routes and the routes to be advertised.

1. Configure the MCE:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter BGP-VPN instance view.	ipv4-family vpn-instance <i>vpn-instance-name</i>	N/A
4. Configure an EBGW peer.	peer { <i>group-name</i> <i>ip-address</i> } as-number <i>as-number</i>	N/A
5. Allow the local AS number to appear in the AS_PATH attribute of a received route, and set the maximum number of times that such case is allowed to appear.	peer { <i>group-name</i> <i>ip-address</i> } allow-as-loop [<i>number</i>]	Optional.
6. Redistribute remote site routes advertised by the PE.	import-route <i>protocol</i> [<i>process-id</i> all-processes] [med <i>med-value</i> route-policy <i>route-policy-name</i>] *	By default, no route redistribution is configured.
7. Configure a filtering policy to filter the routes to be advertised.	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } export [direct isis <i>process-id</i> ospf <i>process-id</i> rip <i>process-id</i> static]	Optional. By default, BGP does not filter the routes to be advertised.
8. Configure a filtering policy to filter the received routes.	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } import	Optional. By default, BGP does not filter the received routes.

BGP checks routing loops by examining AS numbers. When EBGP is used, the MCE advertises routing information carrying the local AS number to the site and then receives routing updates from the site. The routing updates carry the AS number of the MCE, so the MCE discards them to avoid routing loops. To enable the MCE to receive such routes, configure the MCE to allow routing loops.

Routes redistributed from OSPF to BGP on the MCE have their OSPF attributes removed. To enable BGP to distinguish routes redistributed from different OSPF domains, you must enable the redistributed routes to carry the OSPF domain ID by configuring the **domain-id** command in OSPF view. The domain ID is added to BGP VPN routes as an extended community attribute.

BGP runs in a BGP VPN instance in the same way as it runs in a normal network.

2. Configure a VPN site:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Configure the MCE as the EBGP peer.	peer { <i>group-name</i> <i>ip-address</i> } as-number <i>as-number</i>	N/A
4. Redistribute the IGP routes of the VPN.	import-route <i>protocol</i> [<i>process-id</i>] [med <i>med-value</i> route-policy <i>route-policy-name</i>] *	Optional. A VPN site must advertise the VPN network addresses it can reach to the connected MCE.

Configuring IBGP between MCE and VPN site

If IBGP is used for exchanging routing information between an MCE and VPN sites, you must configure a BGP peer for each VPN instance respectively, and redistribute the IGP routes of each VPN instance on the VPN sites.

1. Configure the MCE:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter BGP-VPN instance view.	ipv4-family vpn-instance <i>vpn-instance-name</i>	N/A
4. Configure an IBGP peer.	peer { <i>group-name</i> <i>ip-address</i> } as-number <i>as-number</i>	N/A
5. Configure the system to be the RR and specify the peer as the client of the RR.	peer { <i>group-name</i> <i>ip-address</i> } reflect-client	Optional. By default, no RR or RR client is configured.
6. Redistribute remote site routes advertised by the PE.	import-route <i>protocol</i> [<i>process-id</i>] all-processes] [med <i>med-value</i> route-policy <i>route-policy-name</i>] *	By default, no route redistribution is configured.
7. Configure a filtering policy to filter the routes to be advertised.	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } export [direct isis <i>process-id</i> ospf <i>process-id</i> rip <i>process-id</i> static]	Optional. By default, BGP does not filter the routes to be advertised.

Step	Command	Remarks
8. Configure a filtering policy to filter the received routes.	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } import	Optional. By default, BGP does not filter the received routes.

NOTE:

After you configure a VPN site as an IBGP peer of the MCE, the MCE does not advertise the BGP routes learned from the VPN site to other IBGP peers, including VPNv4 peers. Only when you configure the VPN site as a client of the RR (the MCE), does the MCE advertise routes learned from it to other IBGP peers.

2. Configure a VPN site:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Configure the MCE as the IBGP peer.	peer { <i>group-name</i> <i>ip-address</i> } as-number <i>as-number</i>	N/A
4. Redistribute the IGP routes of the VPN.	import-route <i>protocol</i> [<i>process-id</i>] [med <i>med-value</i> route-policy <i>route-policy-name</i>] *	Optional. A VPN site must advertise the VPN network addresses it can reach to the connected MCE.

Configuring routing between MCE and PE

MCE-PE routing configuration includes these tasks:

- Bind the MCE-PE interfaces to VPN instances
- Perform route configurations
- Redistribute VPN routes into the routing protocol running between the MCE and the PE.

Configuring static routing between MCE and PE

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Configure a static route for a VPN instance.	<ul style="list-style-type: none"> ip route-static <i>dest-address</i> { <i>mask</i> <i>mask-length</i> } { <i>gateway-address</i> <i>interface-type interface-number</i> [<i>gateway-address</i>] vpn-instance <i>d-vpn-instance-name gateway-address</i> } [preference <i>preference-value</i>] [tag <i>tag-value</i>] [description <i>description-text</i>] ip route-static vpn-instance <i>s-vpn-instance-name</i><1-6> <i>dest-address</i> { <i>mask</i> <i>mask-length</i> } { <i>gateway-address</i> [public] <i>interface-type interface-number</i> [<i>gateway-address</i>] vpn-instance <i>d-vpn-instance-name gateway-address</i> } [preference <i>preference-value</i>] [tag <i>tag-value</i>] [description <i>description-text</i>] 	Use either command.
3. Configure the default precedence for static routes.	ip route-static default-preference <i>default-preference-value</i>	Optional. 60 by default.

Configuring RIP between MCE and PE

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a RIP process for a VPN instance and enter RIP view.	rip [<i>process-id</i>] vpn-instance <i>vpn-instance-name</i>	N/A
3. Enable RIP on the interface attached to the specified network.	network <i>network-address</i>	By default, RIP is disabled on an interface.
4. Redistribute the VPN routes.	import-route <i>protocol</i> [<i>process-id</i>] [allow-ibgp] [cost <i>cost</i> route-policy <i>route-policy-name</i> tag <i>tag</i>] *	By default, no route of any other routing protocol is redistributed into RIP.
5. Configure the default cost value for the redistributed routes.	default cost <i>value</i>	Optional. 0 by default.

Configuring OSPF between MCE and PE

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create an OSPF process for a VPN instance and enter OSPF view.	ospf [<i>process-id</i> router-id <i>router-id</i> vpn-instance <i>vpn-instance-name</i>] *	N/A

Step	Command	Remarks
3. Disable routing loop detection.	vpn-instance-capability simple	Disabled by default. You must disable routing loop detection for a VPN OSPF process on the MCE. Otherwise, the MCE cannot receive OSPF routes from the PE.
4. Configure the OSPF domain ID.	domain-id <i>domain-id</i> [secondary]	Optional. 0 by default.
5. Redistribute the VPN routes.	import-route <i>protocol</i> [<i>process-id</i> allow-ibgp] [cost <i>cost</i> type <i>type</i> tag <i>tag</i> route-policy <i>route-policy-name</i>] *	By default, no route of any other routing protocol is redistributed into OSPF.
6. Configure a filtering policy to filter advertised routes.	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } export [<i>protocol</i> [<i>process-id</i>]]	Optional. By default, advertised routes are not filtered.
7. Configure the default parameters for redistributed routes (cost, route number, tag, and type).	default { cost <i>cost</i> limit <i>limit</i> tag <i>tag</i> type <i>type</i> } *	Optional. The default cost is 1, the default maximum number of routes redistributed per time is 1000, the default tag is 1, and default type of redistributed routes is Type-2.
8. Create an OSPF area and enter OSPF area view.	area <i>area-id</i>	By default, no OSPF area is created.
9. Enable OSPF on the interface attached to the specified network in the area.	network <i>ip-address</i> <i>wildcard-mask</i>	By default, an interface neither belongs to any area nor runs OSPF.

Configuring IS-IS between MCE and PE

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create an IS-IS process for a VPN instance and enter IS-IS view.	isis [<i>process-id</i>] vpn-instance <i>vpn-instance-name</i>	N/A
3. Configure a network entity title.	network-entity <i>net</i>	Not configured by default.
4. Redistribute the VPN routes.	import-route { isis [<i>process-id</i>] ospf [<i>process-id</i>] rip [<i>process-id</i>] bgp [allow-ibgp] direct static } [cost <i>cost</i> cost-type { external internal }] [level-1 level-1-2 level-2] route-policy <i>route-policy-name</i> tag <i>tag</i>] *	Optional. By default, IS-IS does not redistribute routes of any other routing protocol. If you do not specify the route level in the command, the command will redistribute routes to the level-2 routing table by default.

Step	Command	Remarks
5. Configure a filtering policy to filter advertised routes.	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> route-policy <i>route-policy-name</i> } export [isis <i>process-id</i> ospf <i>process-id</i> rip <i>process-id</i> bgp direct static]	Optional. By default, IS-IS does not filter advertised routes.
6. Return to system view.	quit	N/A
7. Enter interface view.	interface <i>interface-type interface-number</i>	N/A
8. Enable the IS-IS process on the interface.	isis enable [<i>process-id</i>]	Disabled by default.

Configuring EBGP between MCE and PE

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter BGP-VPN instance view.	ipv4-family vpn-instance <i>vpn-instance-name</i>	N/A
4. Configure the PE as the EBGP peer.	peer { <i>group-name</i> <i>ip-address</i> } as-number <i>as-number</i>	N/A
5. Redistribute the VPN routes of the VPN site.	import-route <i>protocol</i> [<i>process-id</i> all-processes] [med <i>med-value</i> route-policy <i>route-policy-name</i>] *	By default, no route redistribution is configured.
6. Configure a filtering policy to filter the routes to be advertised.	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } export [direct isis <i>process-id</i> ospf <i>process-id</i> rip <i>process-id</i> static]	Optional. By default, BGP does not filter the routes to be advertised.
7. Configure a filtering policy to filter the received routes.	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } import	Optional. By default, BGP does not filter the received routes.

NOTE:

BGP runs within a VPN in the same way as it runs within a public network.

Configuring IBGP between MCE and PE

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter BGP-VPN instance view.	ipv4-family vpn-instance <i>vpn-instance-name</i>	N/A
4. Configure the PE as the IBGP peer.	peer { <i>group-name</i> <i>ip-address</i> } as-number <i>as-number</i>	N/A

Step	Command	Remarks
5. Redistribute the VPN routes of the VPN site.	import-route <i>protocol</i> [<i>process-id</i> all-processes] [med <i>med-value</i> route-policy <i>route-policy-name</i>] *	By default, no route redistribution is configured.
6. Configure the egress router of the site as a client of the route reflector.	peer { <i>group-name</i> <i>ip-address</i> } reflect-client	Optional. By default, no route reflector or client is configured.
7. Enable route reflection between clients.	reflect between-clients	Optional. Enabled by default. If the clients are fully meshed, you do not need to enable route reflection.
8. Specify a cluster ID for the route reflector.	reflector cluster-id <i>cluster-id</i>	Optional. By default, each RR in a cluster uses its own router ID as the cluster ID. If more than one RR exists in a cluster, use this command to configure the same cluster ID for all RRs in the cluster to avoid routing loops.
9. Configure a filtering policy to filter the routes to be advertised.	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } export [direct isis <i>process-id</i> ospf <i>process-id</i> rip <i>process-id</i> static]	Optional. By default, BGP does not filter the routes to be advertised.
10. Configure a filtering policy to filter the received routes.	filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } import	Optional. By default, BGP does not filter the received routes.

Resetting BGP connections

When BGP configuration changes, you can use the soft reset function or reset BGP connections to make new configurations take effect. Soft reset requires that BGP peers have route refreshment capability (supporting Route-Refresh messages).

Step	Command	Remarks
1. Soft reset the BGP connections in a specified VPN instance.	refresh bgp vpn-instance <i>vpn-instance-name</i> { <i>ip-address</i> all external group <i>group-name</i> } { export import }	Available in user view
2. Reset BGP connections of a VPN instance.	reset bgp vpn-instance <i>vpn-instance-name</i> { <i>as-number</i> <i>ip-address</i> all external group <i>group-name</i> }	Available in user view

Displaying and maintaining MCE

Task	Command	Remarks
Display information about the routing table associated with a VPN instance.	display ip routing-table vpn-instance <i>vpn-instance-name</i> [verbose] [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about a specific VPN instance or all VPN instances.	display ip vpn-instance [instance-name <i>vpn-instance-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about the FIB of a VPN instance.	display fib vpn-instance <i>vpn-instance-name</i> [acl <i>acl-number</i> ip-prefix <i>ip-prefix-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about the FIB of a VPN instance that matches the specified destination IP address.	display fib vpn-instance <i>vpn-instance-name</i> <i>ip-address</i> [<i>mask</i> <i>mask-length</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about a specific peer group or all BGP VPNv4 peer groups.	display bgp vpnv4 vpn-instance <i>vpn-instance-name</i> group [<i>group-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about BGP VPNv4 routes injected into a specific VPN instance or all VPN instances.	display bgp vpnv4 vpn-instance <i>vpn-instance-name</i> network [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display BGP VPNv4 AS path information.	display bgp vpnv4 vpn-instance <i>vpn-instance-name</i> paths [<i>as-regular-expression</i> { { begin exclude include } <i>regular-expression</i> }]	Available in any view
Display information about BGP VPNv4 peers.	display bgp vpnv4 vpn-instance <i>vpn-instance-name</i> peer [<i>group-name</i> log-info <i>ip-address</i> { log-info verbose } verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display the BGP VPNv4 routing information of a specific VPN instance.	<pre>display bgp vpnv4 vpn-instance vpn-instance-name routing-table [[network-address [{ mask mask-length } [longer-prefixes]] as-path-acl as-path-acl-number cidr community [aa:nn]&<1-13> [no-advertise no-export no-export-subconfed] * [whole-match] community-list { basic-community-list-number [whole-match] adv-community-list-number }&<1-16> dampened dampening parameter different-origin-as flap-info [network-address [{ mask mask-length } [longer-match]] as-path-acl as-path-acl-number] peer ip-address { advertised-routes received-routes } statistic] [{ begin exclude include } regular-expression] [flap-info] regular-expression as-regular-expression]</pre>	Available in any view
Clear the route flap dampening information of a VPN instance.	<pre>reset bgp vpn-instance vpn-instance-name dampening [network-address [mask mask-length]</pre>	Available in user view
Clear route flap history information about a BGP peer of a VPN instance.	<pre>reset bgp vpn-instance vpn-instance-name ip-address flap-info reset bgp vpn-instance vpn-instance-name flap-info [ip-address [mask mask-length] as-path-acl as-path-acl-number regexp as-path-regexp]</pre>	Available in user view

For commands to display information about a routing table, see *Layer 3—IP Routing Command Reference*.

MCE configuration examples

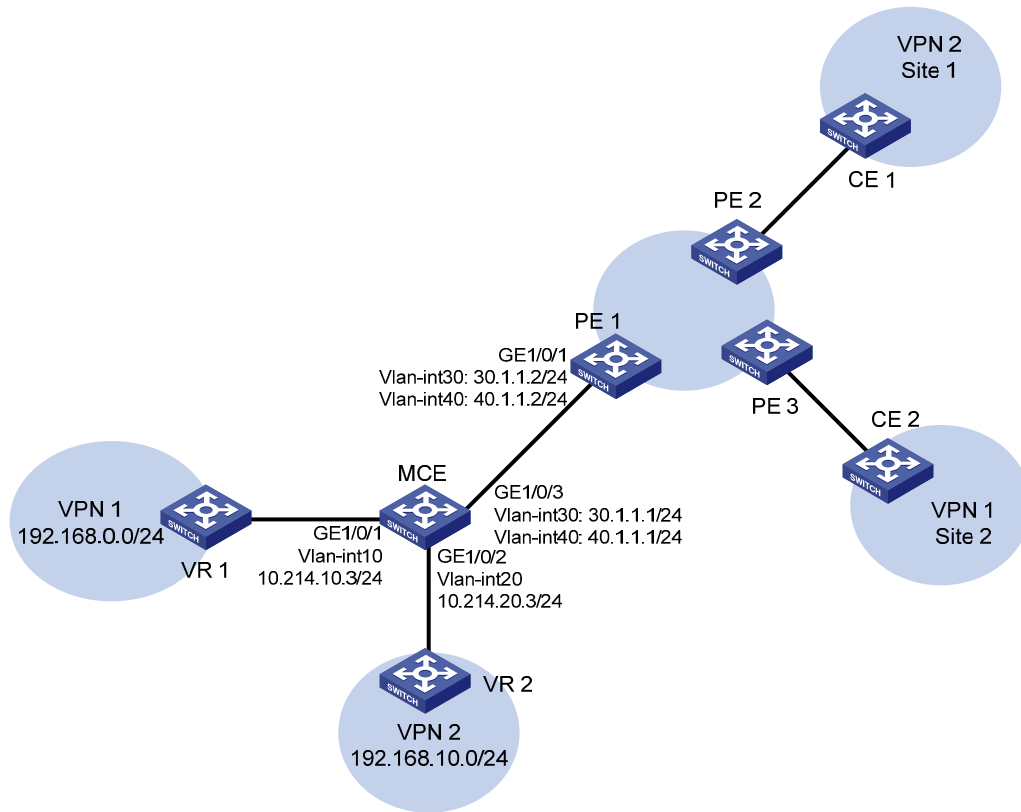
Using OSPF to advertise VPN routes to the PE

Network requirements

As shown in [Figure 129](#), the MCE device is connected to VPN 1 through VLAN-interface 10 and to VPN 2 through VLAN-interface 20. RIP runs in VPN 2.

Configure the MCE device to separate routes from different VPNs and advertise the VPN routes to PE 1 through OSPF.

Figure 129 Network diagram



Configuration procedure

Assume that the system name of the MCE device is MCE, the system names of the edge devices of VPN 1 and VPN 2 are VR1 and VR2, respectively, and the system name of PE 1 is PE1.

1. Configure the VPN instances on the MCE and PE 1:

On the MCE, configure VPN instances **vpn1** and **vpn2**, and specify an RD and route targets for each VPN instance.

```
<MCE> system-view
[MCE] ip vpn-instance vpn1
[MCE-vpn-instance-vpn1] route-distinguisher 10:1
[MCE-vpn-instance-vpn1] vpn-target 10:1
[MCE-vpn-instance-vpn1] quit
[MCE] ip vpn-instance vpn2
[MCE-vpn-instance-vpn2] route-distinguisher 20:1
[MCE-vpn-instance-vpn2] vpn-target 20:1
[MCE-vpn-instance-vpn2] quit
```

Create VLAN 10, add port GigabitEthernet 1/0/1 to VLAN 10, and create VLAN-interface 10.

```
[MCE] vlan 10
[MCE-vlan10] port gigabitethernet 1/0/1
[MCE-vlan10] quit
[MCE] interface vlan-interface 10
```

Bind VLAN-interface 10 with VPN instance **vpn1**, and configure an IP address for VLAN-interface 10.

```
[MCE-vlan-interface10] ip binding vpn-instance vpn1
```

```
[MCE-Vlan-interface10] ip address 10.214.10.3 24
# Configure VLAN 20, add port GigabitEthernet 1/0/2 to VLAN 20, bind VLAN-interface 20 with
VPN instance vpn2, and specify an IP address for VLAN-interface 20.
[MCE-Vlan-interface10] quit
[MCE] vlan 20
[MCE-vlan20] port gigabitethernet 1/0/2
[MCE-vlan20] quit
[MCE] interface vlan-interface 20
[MCE-Vlan-interface20] ip binding vpn-instance vpn2
[MCE-Vlan-interface20] ip address 10.214.20.3 24
[MCE-Vlan-interface20] quit
# On PE 1, configure VPN instances vpn1 and vpn2, specify an RD and route targets for each VPN
instance.
```

```
<PE1> system-view
[PE1] ip vpn-instance vpn1
[PE1-vpn-instance-vpn1] route-distinguisher 30:1
[PE1-vpn-instance-vpn1] vpn-target 10:1
[PE1-vpn-instance-vpn1] quit
[PE1] ip vpn-instance vpn2
[PE1-vpn-instance-vpn2] route-distinguisher 40:1
[PE1-vpn-instance-vpn2] vpn-target 20:1
[PE1-vpn-instance-vpn2] quit
```

2. Configure routing between the MCE and VPN sites:

The MCE is connected to VPN 1 directly, and no routing protocol is enabled in VPN 1. Therefore, you can configure static routes.

```
# On VR 1, assign IP address 10.214.10.2/24 to the interface connected to MCE and
192.168.0.1/24 to the interface connected to VPN 1. Add ports to VLANs correctly. (Details not
shown.)
```

```
# On VR 1, configure a default route with the next hop as 10.214.10.3.
```

```
<VR1> system-view
[VR1] ip route-static 0.0.0.0 0.0.0.0 10.214.10.3
```

```
# On the MCE, configure a static route to 192.168.0.0/24, specify the next hop as 10.214.10.2,
and bind the static route with VPN instance vpn1.
```

```
[MCE] ip route-static vpn-instance vpn1 192.168.0.0 24 10.214.10.2
```

```
# On the MCE, display the routing information maintained for VPN instance vpn1.
```

```
[MCE] display ip routing-table vpn-instance vpn1
Routing Tables: vpn1
          Destinations : 5          Routes : 5
Destination/Mask    Proto  Pre  Cost           NextHop           Interface
10.214.10.0/24      Direct  0    0           10.214.10.3       Vlan10
10.214.10.3/32      Direct  0    0           127.0.0.1         InLoop0
127.0.0.0/8         Direct  0    0           127.0.0.1         InLoop0
127.0.0.1/32       Direct  0    0           127.0.0.1         InLoop0
192.168.0.0/24     Static  60   0           10.214.10.2       Vlan10
```

The output shows that the MCE has a static route for VPN instance **vpn1**.

Run RIP in VPN 2. Create RIP process 20 and bind it with VPN instance **vpn2** on the MCE, so that the MCE can learn the routes of VPN 2 and add them to the routing table of the VPN instance **vpn2**.

```
[MCE] rip 20 vpn-instance vpn2
```

Advertise subnet 10.214.20.0.

```
[MCE-rip-20] network 10.214.20.0
```

```
[MCE-rip-20] quit
```

On VR 2, assign IP address 10.214.20.2/24 to the interface connected to MCE and 192.168.10.1/24 to the interface connected to VPN 2. (Details not shown.)

Configure RIP, and advertise subnets 192.168.10.0 and 10.214.20.0.

```
<VR2> system-view
```

```
[VR2] rip 20
```

```
[VR2-rip-20] network 192.168.10.0
```

```
[VR2-rip-20] network 10.214.20.0
```

On the MCE, display the routing information maintained for VPN instance **vpn2**.

```
[MCE] display ip routing-table vpn-instance vpn2
```

```
Routing Tables: vpn2
```

```
Destinations : 5          Routes : 5
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
10.214.20.0/24	Direct	0	0	10.214.20.3	Vlan20
10.214.20.3/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0
192.168.10.0/24	RIP	100	1	10.214.20.2	Vlan20

The output shows that the MCE has learned the private routes of VPN 2. The MCE maintains the routes of VPN 1 and those of VPN2 in two different routing tables. In this way, routes from different VPNs are separated.

3. Configure routing between MCE and PE 1:

The MCE uses port GigabitEthernet 1/0/3 to connect to PE's port GigabitEthernet 1/0/1. Configure the two ports as trunk ports, and configure them to permit packets carrying VLAN tags 30 and 40 to pass.

```
[MCE] interface gigabitethernet 1/0/3
```

```
[MCE-GigabitEthernet1/0/3] port link-type trunk
```

```
[MCE-GigabitEthernet1/0/3] port trunk permit vlan 30 40
```

```
[MCE-GigabitEthernet1/0/3] quit
```

Configure port GigabitEthernet1/0/1 on the PE.

```
[PE1] interface gigabitethernet 1/0/1
```

```
[PE1-GigabitEthernet1/0/1] port link-type trunk
```

```
[PE1-GigabitEthernet1/0/1] port trunk permit vlan 30 40
```

```
[PE1-GigabitEthernet1/0/1] quit
```

On the MCE, create VLAN 30 and VLAN-interface 30, bind the VLAN interface with VPN instance **vpn1**, and configure an IP address for the VLAN interface.

```
[MCE] vlan 30
```

```
[MCE-vlan30] quit
```

```
[MCE] interface vlan-interface 30
```

```
[MCE-Vlan-interface30] ip binding vpn-instance vpn1
[MCE-Vlan-interface30] ip address 30.1.1.1 24
[MCE-Vlan-interface30] quit
```

On the MCE, create VLAN 40 and VLAN-interface 40, bind the VLAN interface with VPN instance **vpn2**, and configure an IP address for the VLAN interface.

```
[MCE] vlan 40
[MCE-vlan40] quit
[MCE] interface vlan-interface 40
[MCE-Vlan-interface40] ip binding vpn-instance vpn2
[MCE-Vlan-interface40] ip address 40.1.1.1 24
[MCE-Vlan-interface40] quit
```

On PE 1, create VLAN 30 and VLAN-interface 30, bind the VLAN interface with VPN instance **vpn1**, and configure an IP address for the VLAN interface.

```
[PE1] vlan 30
[PE1-vlan30] quit
[PE1] interface vlan-interface 30
[PE1-Vlan-interface30] ip binding vpn-instance vpn1
[PE1-Vlan-interface30] ip address 30.1.1.2 24
[PE1-Vlan-interface30] quit
```

On PE 1, create VLAN 40 and VLAN-interface 40, bind the VLAN interface with VPN instance **vpn2**, and configure an IP address for the VLAN interface.

```
[PE1] vlan 40
[PE1-vlan40] quit
[PE1] interface vlan-interface 40
[PE1-Vlan-interface40] ip binding vpn-instance vpn2
[PE1-Vlan-interface40] ip address 40.1.1.2 24
[PE1-Vlan-interface40] quit
```

Configure the IP address of the interface Loopback0 as 101.101.10.1 for the MCE and as 100.100.10.1 for PE 1. Specify the loopback interface address as the router ID for the MCE and PE 1. (Details not shown.)

Enable OSPF process 10 on the MCE, bind the process to VPN instance **vpn1**, and set the domain ID to 10.

```
[MCE] ospf 10 router-id 101.101.10.1 vpn-instance vpn1
[MCE-ospf-10] vpn-instance-capability simple
[MCE-ospf-10] domain-id 10
```

On the MCE, advertise subnet 30.1.1.0 in area 0, and redistribute the static route of VPN 1.

```
[MCE-ospf-10] area 0
[MCE-ospf-10-area-0.0.0.0] network 30.1.1.0 0.0.0.255
[MCE-ospf-10-area-0.0.0.0] quit
[MCE-ospf-10] import-route static
```

On PE 1, start OSPF process 10, bind the process with VPN instance **vpn1**, set the domain ID to 10, and advertise subnet 30.1.1.0 in area 0.

```
[PE1] ospf 10 router-id 100.100.10.1 vpn-instance vpn1
[PE1-ospf-10] domain-id 10
[PE1-ospf-10] area 0
[PE1-ospf-10-area-0.0.0.0] network 30.1.1.0 0.0.0.255
[PE1-ospf-10-area-0.0.0.0] quit
```

```
[PE1-ospf-10] quit
```

On PE 1, display the routing table of VPN1.

```
[PE1] display ip routing-table vpn-instance vpn1
```

```
Routing Tables: vpn1
```

```
          Destinations : 5          Routes : 5
Destination/Mask    Proto  Pre  Cost           NextHop         Interface
30.1.1.0/24         Direct  0    0             30.1.1.2        Vlan30
30.1.1.2/32         Direct  0    0             127.0.0.1       InLoop0
127.0.0.0/8         Direct  0    0             127.0.0.1       InLoop0
127.0.0.1/32        Direct  0    0             127.0.0.1       InLoop0
192.168.0.0/24      O_ASE  150  1             30.1.1.1        Vlan30
```

The output shows that the static route of VPN 1 has been redistributed to the OSPF routing table of PE 1.

Take similar procedures to configure OSPF process 20 between MCE and PE 1 and redistribute VPN 2's routing information from RIP into the OSPF routing table of MCE. The following output shows that PE 1 has learned the private route of VPN 2 through OSPF.

```
<PE1> display ip routing-table vpn-instance vpn2
```

```
Routing Tables: vpn2
```

```
          Destinations : 5          Routes : 5
Destination/Mask    Proto  Pre  Cost           NextHop         Interface
40.1.1.0/24         Direct  0    0             40.1.1.2        Vlan40
40.1.1.2/32         Direct  0    0             127.0.0.1       InLoop0
127.0.0.0/8         Direct  0    0             127.0.0.1       InLoop0
127.0.0.1/32        Direct  0    0             127.0.0.1       InLoop0
192.168.10.0/24     O_ASE  150  1             40.1.1.1        Vlan40
```

Now, the routing information of the two VPNs has been redistributed into the routing tables on PE 1.

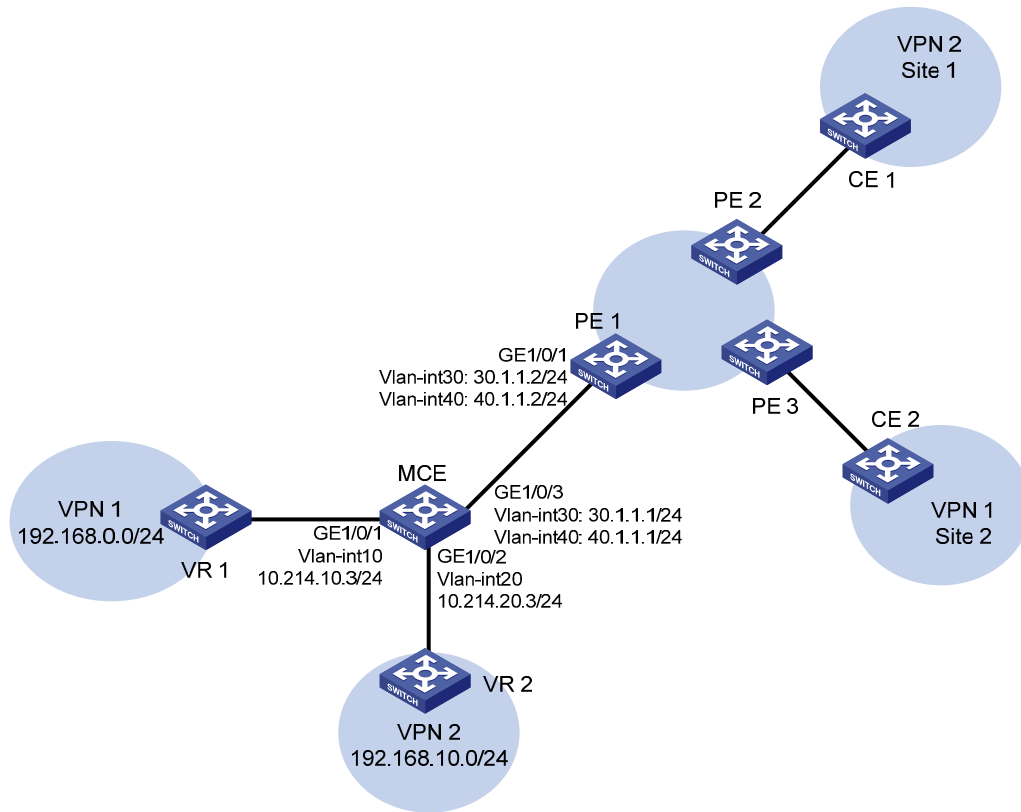
Using BGP to advertise VPN routes to the PE

Network requirements

As shown in [Figure 130](#), use an Ethernet switch as the MCE device. Advertise the VPN routes in site 1 and site 2 to PE 1, so that a VPN's sites across the MPLS backbone network can communicate with each other normally.

Use OSPF in both site 1 and site 2. Use EBGP between the MCE and PE 1.

Figure 130 Network diagram



Configuration procedure

1. Configure VPN instances:
Create VPN instances on the MCE and PE 1, and bind the VPN instances with VLAN interfaces. For the configuration procedure, see "[Using OSPF to advertise VPN routes to the PE.](#)"
2. Configure routing between the MCE and VPN sites:
Start an OSPF process on the devices in the two VPNs and advertise the subnets. (Details not shown.)
Configure OSPF on the MCE, and bind OSPF process 10 with VPN instance **vpn1** to learn the routes of VPN 1.

```
<MCE> system-view
[MCE] ospf router-id 10.214.10.3 10 vpn-instance vpn1
[MCE-ospf-10] area 0
[MCE-ospf-10-area-0.0.0.0] network 10.214.10.0 0.0.0.255
```

Display the routing table of VPN 1 on the MCE.

```
[MCE-ospf-10-area-0.0.0.0] display ip routing-table vpn-instance vpn1
```

Routing Tables: vpn1

Destinations : 5 Routes : 5

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
10.214.10.0/24	Direct	0	0	10.214.10.3	Vlan10
10.214.10.3/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0


```

127.0.0.1/32      Direct 0    0                127.0.0.1      InLoop0
192.168.0.0/24   OSPF   10    1                10.214.10.2    Vlan10

```

The output shows that the MCE has learned the private route of VPN 1 through OSPF process 10.

On MCE, bind OSPF process 20 with VPN instance **vpn2** to learn the routes of VPN 2. The configuration procedure is similar to that for OSPF process 10.

The following output shows that the MCE has learned the private route of VPN 2 through OSPF:

```
[MCE] display ip routing-table vpn-instance vpn2
```

```
Routing Tables: vpn2
```

```
Destinations : 5          Routes : 5
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
10.214.20.0/24	Direct	0	0	10.214.20.3	Vlan20
10.214.20.3/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0
192.168.0.0/24	OSPF	10	1	10.214.20.2	Vlan20

3. Configure routing between the MCE and PE 1:

Configure the connecting ports between the MCE and PE 1 as trunk ports. The configuration procedure is similar to that described in "[Using OSPF to advertise VPN routes to the PE.](#)" (Details not shown.)

Start BGP process 100 on the MCE, and enter the IPv4 address family view of VPN instance **vpn1**.

```
[MCE] bgp 100
```

```
[MCE-bgp] ipv4-family vpn-instance vpn1
```

Specify PE 1 as the EBGP peer of the MCE, and redistribute the routing information of OSPF process 10. (The IP address of PE 1's interface bound with VPN instance **vpn1** is 10.100.10.3, and the BGP process is 200.)

```
[MCE-bgp-vpn1] peer 30.1.1.2 as-number 200
```

```
[MCE-BGP-vpn1] import-route ospf 10
```

On PE 1, configure BGP process 200 and specify the MCE as its EBGP peer.

```
<PE1> system-view
```

```
[PE1] bgp 200
```

```
[PE1-bgp] ipv4-family vpn-instance vpn1
```

```
[PE1-bgp-vpn1] peer 30.1.1.1 as-number 100
```

```
[PE1-bgp-vpn1] quit
```

```
[PE1-bgp] quit
```

On PE 1, display the routing information of VPN instance **vpn1**.

```
[PE1] display ip routing-table vpn-instance vpn1
```

```
Routing Tables: vpn1
```

```
Destinations : 5          Routes : 5
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
30.1.1.0/24	Direct	0	0	30.1.1.2	Vlan30
30.1.1.2/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0

```

127.0.0.1/32      Direct 0    0          127.0.0.1      InLoop0
192.168.0.0/24   BGP     255  2          30.1.1.1       Vlan30

```

Perform similar configuration on the MCE and PE 1 for VPN 2. Redistribute the OSPF routes of VPN instance **vpn2** into the EBGp routing table. (Details not shown.)

The following output shows that PE 1 has learned the private route of VPN 2 through BGP:

```
[PE1] display ip routing-table vpn-instance vpn2
```

```
Routing Tables: vpn2
```

```
Destinations : 5      Routes : 5
```

Destination/Mask	Proto	Pre	Cost	NextHop	Interface
40.1.1.0/24	Direct	0	0	40.1.1.2	Vlan40
40.1.1.2/32	Direct	0	0	127.0.0.1	InLoop0
127.0.0.0/8	Direct	0	0	127.0.0.1	InLoop0
127.0.0.1/32	Direct	0	0	127.0.0.1	InLoop0
192.168.10.0/24	BGP	255	2	40.1.1.1	Vlan40

Now, the MCE has redistributed the OSPF routes of the two VPN instances into the EBGp routing tables of PE 1.

Configuring IPv6 MCE

The IPv6 MCE function is available only on the HP 5500 EI switch series.

Overview

In an IPv6 MPLS L3 VPN, an IPv6 MCE advertises IPv6 routing information between the VPN and the connected PE and forwards IPv6 packets. An IPv6 MCE operates in the same way as an IPv4 MCE. For more information, see "Configuring MCE."

Configuring an IPv6 MCE

Configuring VPN instances

By configuring VPN instances on a PE, you isolate not only VPN routes from public network routes, but also routes of a VPN from those of another VPN. This feature allows VPN instances to be used in network scenarios besides MPLS L3VPNs.

Creating a VPN instance

A VPN instance is associated with a site. It is a collection of the VPN membership and routing rules of its associated site. A VPN instance does not necessarily correspond to one VPN.

A VPN instance takes effect only after you configure an RD for it.

You can configure a description for a VPN instance to record its related information, such as its relationship with a certain VPN.

To create and configure a VPN instance:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a VPN instance and enter VPN instance view.	ip vpn-instance <i>vpn-instance-name</i>	N/A
3. Configure an RD for the VPN instance.	route-distinguisher <i>route-distinguisher</i>	N/A
4. Configure a description for the VPN instance.	description <i>text</i>	Optional.

Associating a VPN instance with an interface

After creating and configuring a VPN instance, you need to associate the VPN instance with the interface for connecting the CE. Any LDP-capable interface can be associated with a VPN instance.

To associate a VPN instance with an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Associate a VPN instance with the interface.	ip binding vpn-instance <i>vpn-instance-name</i>	No VPN instance is associated with an interface by default.

NOTE:

The **ip binding vpn-instance** command clears the IPv6 address of the interface on which it is configured. Be sure to re-configure an IPv6 address for the interface after configuring the command.

Configuring route related attributes for a VPN instance

The control process of VPN route advertisement is as follows:

- When a VPN route learned from a CE gets redistributed into BGP, BGP associates it with a route target extended community attribute list, which is usually the export target attribute of the VPN instance associated with the CE.
- The VPN instance determines which routes it can accept and redistribute according to the **import-extcommunity** in the route target.
- The VPN instance determines how to change the route targets attributes for routes to be advertised according to the **export-extcommunity** in the route target.

! IMPORTANT:

Create a routing policy before associating it with a VPN instance. Otherwise, the switch cannot filter the routes to be received and advertised.

To configure route related attributes for a VPN instance:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VPN instance view.	ip vpn-instance <i>vpn-instance-name</i>	N/A
3. Enter IPv6 VPN view.	ipv6-family	Optional.
4. Configure route targets.	vpn-target <i>vpn-target</i> <1-8> [both export-extcommunity import-extcommunity]	A single vpn-target command can configure up to eight VPN targets. You can configure up to 64 VPN targets for a VPN instance.

Step	Command	Remarks
5. Set the maximum number of routes supported.	routing-table limit <i>number</i> { <i>warn-threshold</i> simply-alert }	Optional. Setting the maximum number of routes for a VPN instance to support is for preventing too many routes from being redistributed into the PE.
6. Apply an import routing policy.	import route-policy <i>route-policy</i>	Optional. By default, all routes matching the import target attribute are accepted.
7. Apply an export routing policy.	export route-policy <i>route-policy</i>	Optional. By default, routes to be advertised are not filtered.

NOTE:

- Route related attributes configured in VPN instance view are applicable to both IPv4 VPNs and IPv6 VPNs.
- You can configure route related attributes for IPv6 VPNs in both VPN instance view and IPv6 VPN view. Those configured in IPv6 VPN view take precedence.

Configuring routing on an IPv6 MCE

An IPv6 MCE implements service isolation through route isolation. IPv6 MCE routing configuration includes:

- IPv6 MCE-VPN site routing configuration
- IPv6 MCE-PE routing configuration

On the PE in an IPv6 MCE network environment, disable routing loop detection to avoid route loss during route calculation and disable route redistribution between routing protocols to save system resources.

Configuration prerequisites

Before you configure routing on an IPv6 MCE, complete the following tasks:

- On the IPv6 MCE, configure VPN instances, and bind the VPN instances with the interfaces connected to the VPN sites and those connected to the PE.
- Configure the link layer and network layer protocols on related interfaces to ensure IP connectivity.

Configuring routing between IPv6 MCE and VPN site

Configuring static routing between IPv6 MCE and VPN site

An IPv6 MCE can reach a VPN site through an IPv6 static route. IPv6 static routing on a traditional CE is globally effective and thus does not support address overlapping among VPNs. An IPv6 MCE supports

binding an IPv6 static route with an IPv6 VPN instance, so that the IPv6 static routes of different IPv6 VPN instances can be isolated from each other.

To configure IPv6 static routing between IPv6 MCE and VPN site:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure an IPv6 static route for an IPv6 VPN instance.	<ul style="list-style-type: none"> ipv6 route-static <i>ipv6-address prefix-length</i> { <i>interface-type interface-number</i> [<i>next-hop-address</i>] <i>next-hop-address</i> vpn-instance <i>d-vpn-instance-name nexthop-address</i> } [preference <i>preference-value</i>] ipv6 route-static vpn-instance <i>s-vpn-instance-name</i> <1-6> <i>ipv6-address prefix-length</i> { <i>interface-type interface-number</i> [<i>next-hop-address</i>] <i>next-hop-address</i> [public] vpn-instance <i>d-vpn-instance-name nexthop-address</i> } [preference <i>preference-value</i>] 	<p>Use either command.</p> <p>Perform this configuration on the IPv6 MCE. On a VPN site, configure normal IPv6 static routes.</p>

Configuring RIPng between IPv6 MCE and VPN site

A RIPng process belongs to the public network or a single IPv6 VPN instance. If you create a RIPng process without binding it to an IPv6 VPN instance, the process belongs to the public network. By configuring RIPng process-to-IPv6 VPN instance bindings on an IPv6 MCE, you allow routes of different VPNs to be exchanged between the IPv6 MCE and the sites through different RIPng processes, ensuring the separation and security of IPv6 VPN routes.

To configure RIPng between IPv6 MCE and VPN site:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a RIPng process for a VPN instance and enter RIPng view.	ripng [<i>process-id</i>] vpn-instance <i>vpn-instance-name</i>	Perform this configuration on the IPv6 MCE. On a VPN site, configure normal RIPng.
3. Redistribute remote site routes advertised by the PE.	import-route <i>protocol</i> [<i>process-id</i>] [allow-ibgp] [cost <i>cost</i>] route-policy <i>route-policy-name</i>] *	By default, no route of any other routing protocol is redistributed into RIPng.
4. Configure the default cost value for the redistributed routes.	default cost <i>value</i>	Optional. 0 by default.
5. Return to system view.	quit	N/A
6. Enter interface view.	interface <i>interface-type interface-number</i>	N/A
7. Enable RIPng on the interface.	ripng <i>process-id</i> enable	Disabled by default.

Configuring OSPFv3 between IPv6 MCE and VPN site

An OSPFv3 process belongs to the public network or a single IPv6 VPN instance. If you create an OSPFv3 process without binding it to an IPv6 VPN instance, the process belongs to the public network.

By configuring OSPFv3 process-to-IPv6 VPN instance bindings on an IPv6 MCE, you allow routes of different IPv6 VPNs to be exchanged between the IPv6 MCE and the sites through different OSPFv3 processes, ensuring the separation and security of IPv6 VPN routes.

To configure OSPFv3 between IPv6 MCE and VPN site:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create an OSPFv3 process for a VPN instance and enter OSPFv3 view.	ospfv3 [<i>process-id</i>] vpn-instance <i>vpn-instance-name</i>	Perform this configuration on the IPv6 MCE. On a VPN site, configure normal OSPFv3.
3. Set the router ID.	router-id <i>router-id</i>	N/A
4. Redistribute remote site routes advertised by the PE..	import-route <i>protocol</i> [<i>process-id</i> allow-ibgp] [cost <i>value</i> route-policy <i>route-policy-name</i> type <i>type</i>] *	By default, no route of any other routing protocol is redistributed into OSPFv3.
5. Return to system view.	quit	N/A
6. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
7. Enable OSPFv3 on the interface.	ospfv3 <i>process-id</i> area <i>area-id</i> [instance <i>instance-id</i>]	By default, OSPFv3 is disabled on an interface.

NOTE:

Deleting a VPN instance also deletes all related OSPFv3 processes at the same time.

Configuring IPv6 IS-IS between IPv6 MCE and VPN site

An IPv6 IS-IS process belongs to the public network or a single IPv6 VPN instance. If you create an IPv6 IS-IS process without binding it to an IPv6 VPN instance, the process belongs to the public network.

By configuring IPv6 IS-IS process-to-IPv6 VPN instance bindings on an IPv6 MCE, you allow routes of different IPv6 VPNs to be exchanged between the IPv6 MCE and the sites through different IPv6 IS-IS processes, ensuring the separation and security of IPv6 VPN routes.

To configure IPv6 IS-IS between IPv6 MCE and VPN site:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create an IPv6 IS-IS process for a VPN instance and enter IS-IS view.	isis [<i>process-id</i>] vpn-instance <i>vpn-instance-name</i>	Perform this configuration on the IPv6 MCE. On a VPN site, configure normal IPv6 IS-IS.
3. Configure a network entity title for the IS-IS process.	network-entity <i>net</i>	Not configured by default.

Step	Command	Remarks
4. Enable the IPv6 capacity for the IPv6 IS-IS process.	ipv6 enable	Disabled by default.
5. Redistribute remote site routes advertised by the PE.	ipv6 import-route <i>protocol</i> [<i>process-id</i>] [allow-ibgp] [cost cost [level-1 level-1-2 level-2] route-policy route-policy-name tag tag] *	Optional. By default, no routes from any other routing protocol are redistributed to IPv6 IS-IS. If you do not specify the route level in the command, redistributed routes are added to the level-2 routing table by default.
6. Return to system view.	quit	N/A
7. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
8. Enable the IPv6 IS-IS process on the interface.	isis ipv6 enable [<i>process-id</i>]	Disabled by default.

Configuring EBGp between IPv6 MCE and VPN site

To use EBGp for exchanging routing information between an IPv6 MCE and IPv6 VPN sites, you must configure a BGP peer for each IPv6 VPN instance on the IPv6 MCE, and redistribute the IGP routes of each VPN instance on the IPv6 VPN sites.

If EBGp is used for route exchange, you also can configure filtering policies to filter the received routes and the routes to be advertised.

1. Configure the IPv6 MCE:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 BGP-VPN instance view.	ipv6-family vpn-instance <i>vpn-instance-name</i>	N/A
4. Specify an IPv6 BGP peer in an AS.	peer <i>ipv6-address</i> as-number <i>as-number</i>	N/A
5. Redistribute remote site routes advertised by the PE.	import-route <i>protocol</i> [<i>process-id</i>] [med med-value route-policy route-policy-name] *	By default, No route redistribution is configured.
6. Configure a filtering policy to filter the routes to be advertised.	filter-policy { <i>acl6-number</i> ipv6-prefix <i>ip-prefix-name</i> } export [direct isisv6 process-id ripng process-id static]	Optional. By default, the IPv6 MCE does not filter the routes to be advertised.
7. Configure a filtering policy to filter the received routes.	filter-policy { <i>acl6-number</i> ipv6-prefix <i>ip-prefix-name</i> } import	Optional. By default, the IPv6 MCE does not filter the received routes.

NOTE:

After you configure an IPv6 BGP VPN instance, the IPv6 BGP route exchange for the IPv6 VPN instance is the same with the normal IPv6 BGP VPN route exchange.

2. Configure a VPN site:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Configure the IPv6 MCE as the EBGp peer.	peer <i>ipv6-address as-number as-number</i>	N/A
5. Redistribute the IGP routes of the VPN.	import-route <i>protocol [process-id [med med-value route-policy route-policy-name] *]</i>	Optional. By default, no route redistribution is configured. A VPN site must advertise the IPv6 VPN network addresses it can reach to the connected IPv6 MCE.

Configuring routing between IPv6 MCE and PE

IPv6 MCE-PE routing configuration includes these tasks:

- Bind the IPv6 MCE-PE interfaces to IPv6 VPN instances
- Perform routing configurations
- Redistribute IPv6 VPN routes into the routing protocol running between the IPv6 MCE and the PE.

Configuring IPv6 static routing between IPv6 MCE and PE

To configure IPv6 static routing between IPv6 MCE and PE:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure an IPv6 static route for an IPv6 VPN instance.	<ul style="list-style-type: none">• ipv6 route-static <i>ipv6-address prefix-length { interface-type interface-number [next-hop-address] next-hop-address vpn-instance d-vpn-instance-name nexthop-address } [preference preference-value]</i>• ipv6 route-static vpn-instance <i>s-vpn-instance-name&<1-6> ipv6-address prefix-length { interface-type interface-number [next-hop-address] nexthop-address [public] vpn-instance d-vpn-instance-name nexthop-address } [preference preference-value]</i>	User either command.

Configuring RIPng between IPv6 MCE and PE

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a RIPng process for an IPv6 VPN instance and enter RIPng view.	ripng [<i>process-id</i>] vpn-instance <i>vpn-instance-name</i>	N/A
3. Redistribute the VPN routes.	import-route <i>protocol</i> [<i>process-id</i>] [allow-ibgp] [cost <i>cost</i> route-policy <i>route-policy-name</i>] *	By default, no route of any other routing protocol is redistributed into RIPng.
4. Configure the default cost value for the redistributed routes.	default cost <i>value</i>	Optional. 0 by default.
5. Return to system view.	quit	N/A
6. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
7. Enable the RIPng process on the interface.	ripng <i>process-id</i> enable	Disabled by default.

Configuring OSPFv3 between IPv6 MCE and PE

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create an OSPFv3 process for an IPv6 VPN instance and enter OSPFv3 view.	ospfv3 [<i>process-id</i>] vpn-instance <i>vpn-instance-name</i>	N/A
3. Set the router ID.	router-id <i>router-id</i>	N/A
4. Redistribute the VPN routes.	import-route <i>protocol</i> [<i>process-id</i> allow-ibgp] [cost <i>value</i> route-policy <i>route-policy-name</i> type <i>type</i>] *	By default, no route of any other routing protocol is redistributed into OSPFv3.
5. Configure a filtering policy to filter the redistributed routes.	filter-policy { <i>acl6-number</i> ipv6-prefix <i>ipv6-prefix-name</i> } export [bgp4+ direct isisv6 <i>process-id</i> ospfv3 <i>process-id</i> ripng <i>process-id</i> static]	Optional. By default, redistributed routes are not filtered.
6. Return to system view.	quit	N/A
7. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
8. Enable the OSPFv3 process on the interface.	ospfv3 <i>process-id</i> area <i>area-id</i> [instance <i>instance-id</i>]	Disabled by default.

Configuring IPv6 IS-IS between IPv6 MCE and PE

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create an IS-IS process for an IPv6 VPN instance and enter IS-IS view.	isis [<i>process-id</i>] vpn-instance <i>vpn-instance-name</i>	N/A
3. Configure a network entity title.	network-entity <i>net</i>	Not configured by default.
4. Enable the IPv6 capacity for the IS-IS process.	ipv6 enable	Disabled by default.
5. Redistribute the VPN routes.	ipv6 import-route <i>protocol</i> [<i>process-id</i>] [allow-ibgp] [cost <i>cost</i> [level-1 level-1-2 level-2] route-policy <i>route-policy-name</i> tag <i>tag</i>] *	Optional. By default, IS-IS does not redistribute routes of any other routing protocol. If you do not specify the route level in the command, the command will redistribute routes to the level-2 routing table by default.
6. Configure a filtering policy to filter the redistributed routes.	ipv6 filter-policy { <i>acl6-number</i> ipv6-prefix <i>ipv6-prefix-name</i> route-policy <i>route-policy-name</i> } export [<i>protocol</i> [<i>process-id</i>]]	Optional. By default, IPv6 IS-IS does not filter redistributed routes.
7. Return to system view.	quit	N/A
8. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
9. Enable IPv6 for the IS-IS process on the interface.	isis ipv6 enable [<i>process-id</i>]	Disabled by default.

Configuring EBGW between IPv6 MCE and PE

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 BGP-VPN instance view.	ipv6-family vpn-instance <i>vpn-instance-name</i>	N/A
4. Configure the PE as the EBGW peer.	peer <i>ipv6-address</i> as-number <i>as-number</i>	N/A
5. Redistribute the VPN routes.	import-route <i>protocol</i> [<i>process-id</i> [med <i>med-value</i> route-policy <i>route-policy-name</i>] *]	By default, No route redistribution is configured.
6. Configure a filtering policy to filter the routes to be advertised.	filter-policy { <i>acl6-number</i> ipv6-prefix <i>ip-prefix-name</i> } export [direct isisv6 <i>process-id</i> ripng <i>process-id</i> static]	Optional. By default, BGP does not filter the routes to be advertised.
7. Configure a filtering policy to filter the received routes.	filter-policy { <i>acl6-number</i> ipv6-prefix <i>ip-prefix-name</i> } import	Optional. By default, BGP does not filter the received routes.

NOTE:

IPv6 BGP runs within a VPN in the same way as it runs within a public network.

Resetting BGP connections

When BGP configuration changes, you can use the soft reset function or reset BGP connections to make new configurations take effect. Soft reset requires that BGP peers have route refreshment capability (supporting Route-Refresh messages).

Use the following commands to hard reset or soft reset BGP connections:

Step	Command	Remarks
1. Soft reset the IPv6 BGP connections in a VPN instance.	refresh bgp ipv6 vpn-instance <i>vpn-instance-name</i> { <i>ipv6-address</i> all external } { export import }	Available in user view
2. Hard reset the IPv6 BGP connections of a VPN instance.	reset bgp ipv6 vpn-instance <i>vpn-instance-name</i> { <i>as-number</i> <i>ipv6-address</i> all external }	Available in user view

Displaying information about IPv6 MCE

Task	Command	Remarks
Display information about a specific or all VPN instances.	display ip vpn-instance [<i>instance-name</i> <i>vpn-instance-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about the IPv6 FIB of a VPN instance.	display ipv6 fib vpn-instance <i>vpn-instance-name</i> [acl6 <i>acl6-number</i> ipv6-prefix <i>ipv6-prefix-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display a VPN instance's FIB entries that match the specified destination IPv6 address.	display ipv6 fib vpn-instance <i>vpn-instance-name</i> <i>ipv6-address</i> [<i>prefix-length</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about BGP VPNv6 peers established between PEs.	display bgp vpnv6 all peer [<i>ipv4-address</i> verbose verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about IPv6 BGP peers established between the PE and CE in a VPN instance.	display bgp vpnv6 vpn-instance <i>vpn-instance-name</i> peer [<i>ipv6-address</i> verbose verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display all BGP VPNv6 routing information.	display bgp vpnv6 all routing-table [<i>network-address</i> <i>prefix-length</i> [longer-prefixes] peer <i>ip-address</i> { advertised-routes received-routes } [statistic] statistic] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display the BGP VPNv6 routing information of a specified RD.	display bgp vpnv6 route-distinguisher <i>route-distinguisher</i> routing-table [<i>network-address prefix-length</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the BGP VPNv6 routing information of a VPN instance.	display bgp vpnv6 vpn-instance <i>vpn-instance-name</i> routing-table [<i>network-address prefix-length</i> [longer-prefixes] peer <i>ipv6-address</i> { advertised-routes received-routes }] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

For commands that display information about a routing table, see *Layer 3—IP Routing Command Reference*.

IPv6 MCE configuration examples

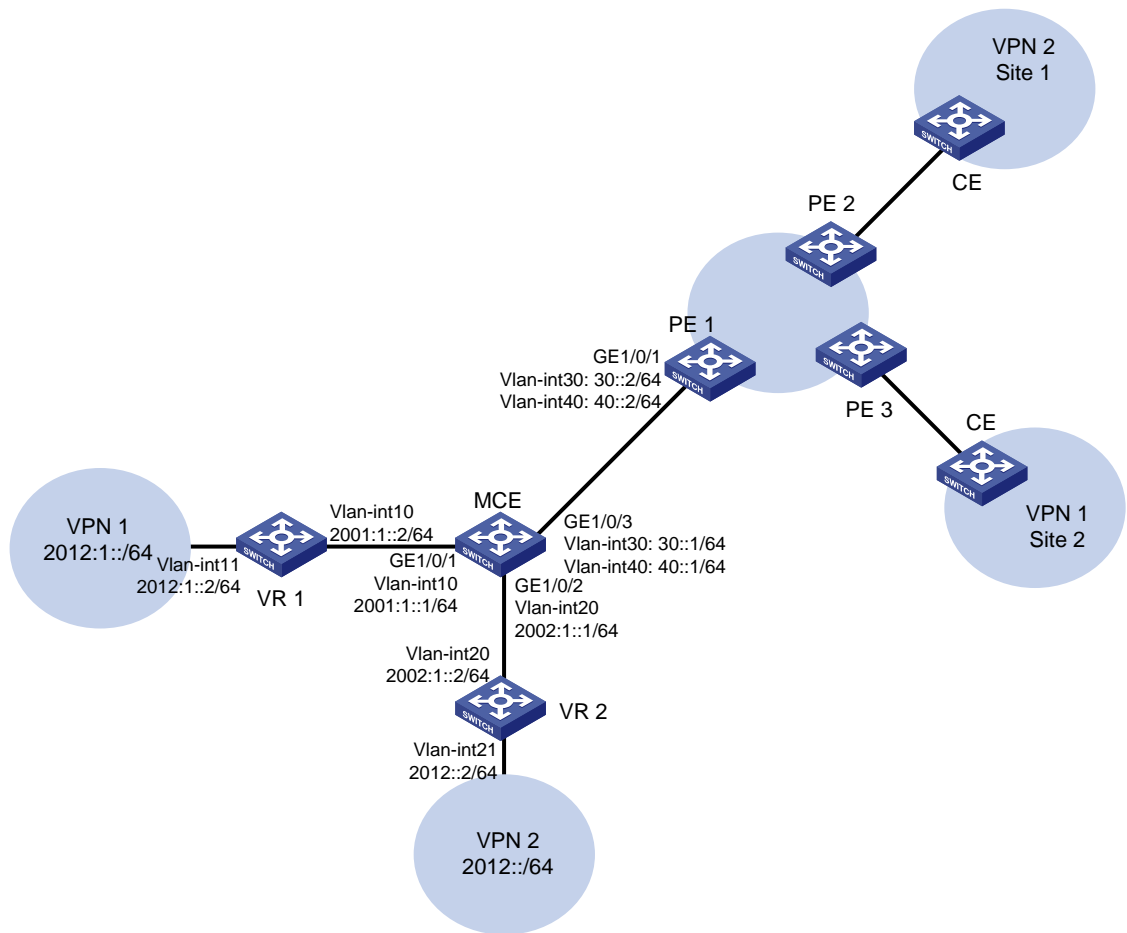
Using IPv6 ISIS to advertise VPN routes to the PE

Network requirements

As shown in [Figure 131](#), the IPv6 MCE device is connected to VPN 1 through VLAN-interface 10 and to VPN 2 through VLAN-interface 20. RIPng is used in VPN 2.

Configure the IPv6 MCE to separate routes from different VPNs and advertise VPN routes to PE 1 through OSPFv3.

Figure 131 Network diagram



Configuration procedure

Assume that the system name of the IPv6 MCE device is MCE, the system names of the edge devices of VPN 1 and VPN 2 are VR1 and VR2, respectively, and the system name of PE 1 is PE1.

1. Configure the VPN instances on the MCE and PE 1.

On the MCE, configure VPN instances **vpn1** and **vpn2**, and specify a RD and route targets for each VPN instance.

```
<MCE> system-view
[MCE] ip vpn-instance vpn1
[MCE-vpn-instance-vpn1] route-distinguisher 10:1
[MCE-vpn-instance-vpn1] vpn-target 10:1
[MCE-vpn-instance-vpn1] quit
[MCE] ip vpn-instance vpn2
[MCE-vpn-instance-vpn2] route-distinguisher 20:1
[MCE-vpn-instance-vpn2] vpn-target 20:1
[MCE-vpn-instance-vpn2] quit
```

Create VLAN 10, add port GigabitEthernet 1/0/1 to VLAN 10, and create VLAN-interface 10.

```
[MCE] vlan 10
[MCE-vlan10] port gigabitethernet 1/0/1
```

```
[MCE-vlan10] quit
```

Bind VLAN-interface 10 with VPN instance **vpn1**, and configure an IPv6 address for the VLAN interface.

```
[MCE] interface vlan-interface 10
[MCE-Vlan-interface10] ip binding vpn-instance vpn1
[MCE-Vlan-interface10] ipv6 address 2001:1::1 64
[MCE-Vlan-interface10] quit
```

Configure VLAN 20, add port GigabitEthernet 1/0/2 to VLAN 20, bind VLAN-interface 20 with VPN instance **vpn2**, and assign an IPv6 address to VLAN-interface 20.

```
[MCE] vlan 20
[MCE-vlan20] port gigabitethernet 1/0/2
[MCE-vlan20] quit
[MCE] interface vlan-interface 20
[MCE-Vlan-interface20] ip binding vpn-instance vpn2
[MCE-Vlan-interface20] ipv6 address 2002:1::1 64
[MCE-Vlan-interface20] quit
```

On PE 1, configure VPN instances **vpn1** and **vpn2**, and specify an RD and route targets for each VPN instance.

```
<PE1> system-view
[PE1] ip vpn-instance vpn1
[PE1-vpn-instance-vpn1] route-distinguisher 30:1
[PE1-vpn-instance-vpn1] vpn-target 10:1
[PE1-vpn-instance-vpn1] quit
[PE1] ip vpn-instance vpn2
[PE1-vpn-instance-vpn2] route-distinguisher 40:1
[PE1-vpn-instance-vpn2] vpn-target 20:1
[PE1-vpn-instance-vpn2] quit
```

2. Configure routing between the MCE and VPN sites.

The MCE is connected with VPN 1 directly, and no routing protocol is enabled in VPN 1. Therefore, you can configure IPv6 static routes.

On VR 1, assign IP address 2001:1::2/64 to the interface connected to the MCE and 2012:1::2/64 to the interface connected to VPN 1. Add ports to VLANs. (Details not shown.)

On VR 1, configure a default route, specifying the next hop as 2001:1::1.

```
<VR1> system-view
[VR1] ipv6 route-static :: 0 2001:1::1
```

On the MCE, configure an IPv6 static route to 2012:1::/64, specify the next hop as 2001:1::2, and bind the static route with VPN instance **vpn1**.

```
[MCE] ipv6 route-static vpn-instance vpn1 2012:1:: 64 vpn-instance vpn1 2001:1::2
```

Run RIPng in VPN 2. Configure RIPng process 20 for VPN instance **vpn2** on the MCE, so that the MCE can learn the routes of VPN 2 and add them to the routing table of VPN instance **vpn2**.

Configure RIPng process 20, binding it with VPN instance **vpn2**.

```
[MCE] ripng 20 vpn-instance vpn2
```

Advertise subnet 2002:1::/64 through RIPng.

```
[MCE] interface vlan-interface 20
[MCE-Vlan-interface20] ripng 20 enable
[MCE-Vlan-interface20] quit
```

On VR 2, assign IPv6 address 2002:1::2/64 to the interface connected to the MCE and 2012::2/64 to the interface connected to VPN 2. (Details not shown.)

Configure RIPng, and advertise subnets 2012::/64 and 2002:1::/64.

```
<VR2> system-view
[VR2] ripng 20
[VR2-ripng-20] quit
[VR2] interface vlan-interface 20
[VR2-Vlan-interface20] ripng 20 enable
[VR2-Vlan-interface20] quit
[VR2] interface vlan-interface 21
[VR2-Vlan-interface21] ripng 20 enable
[VR2-Vlan-interface21] quit
```

On MCE, display the routing tables of VPN instances **vpn1** and **vpn2**.

```
[MCE] display ipv6 routing-table vpn-instance vpn1
```

```
Routing Table : vpn1
Destinations : 5          Routes : 5
```

```
Destination: ::1/128          Protocol : Direct
NextHop      : ::1           Preference: 0
Interface   : InLoop0       Cost      : 0
```

```
Destination: 2001:1::/64     Protocol : Direct
NextHop      : 2001:1::1     Preference: 0
Interface   : Vlan10        Cost      : 0
```

```
Destination: 2001:1::1/128   Protocol : Direct
NextHop      : ::1           Preference: 0
Interface   : InLoop0       Cost      : 0
```

```
Destination: 2012:1::/64     Protocol : Static
NextHop      : 2001:1::2     Preference: 60
Interface   : Vlan10        Cost      : 0
```

```
Destination: FE80::/10       Protocol : Direct
NextHop      : ::            Preference: 0
Interface   : NULL0         Cost      : 0
```

```
[MCE] display ipv6 routing-table vpn-instance vpn2
```

```
Routing Table : vpn2
Destinations : 5          Routes : 5
```

```
Destination: ::1/128          Protocol : Direct
NextHop      : ::1           Preference: 0
Interface   : InLoop0       Cost      : 0
```



```

Destination: 2002:1::/64                                Protocol : Direct
NextHop      : 2002:1::1                                Preference: 0
Interface    : Vlan20                                    Cost      : 0

```

```

Destination: 2002:1::1/128                              Protocol : Direct
NextHop      : ::1                                       Preference: 0
Interface    : InLoop0                                    Cost      : 0
Destination: 2012::/64                                  Protocol : RIPng
NextHop      : FE80::20F:E2FF:FE3E:9CA2                 Preference: 100
Interface    : Vlan20                                    Cost      : 1

```

```

Destination: FE80::/10                                   Protocol : Direct
NextHop      : ::                                         Preference: 0
Interface    : NULL0                                     Cost      : 0

```

The output shows that the MCE has learned the private route of VPN 2. The MCE maintains the routes of VPN 1 and VPN 2 in two different routing tables. In this way, routes from different VPNs are separated.

3. Configure routing between the MCE and PE 1.

On the MCE, configure the port connected to PE 1 as a trunk port, and configure it to permit packets of VLAN 30 and VLAN 40 to pass with VLAN tags.

```

[MCE] interface gigabitethernet 1/0/3
[MCE-GigabitEthernet1/0/3] port link-type trunk
[MCE-GigabitEthernet1/0/3] port trunk permit vlan 30 40
[MCE-GigabitEthernet1/0/3] quit

```

On PE 1, configure the port connected to MCE as a trunk port, and configure it to permit packets of VLAN 30 and VLAN 40 to pass with VLAN tags.

```

[PE1] interface gigabitethernet 1/0/1
[PE1-GigabitEthernet1/0/1] port link-type trunk
[PE1-GigabitEthernet1/0/1] port trunk permit vlan 30 40
[PE1-GigabitEthernet1/0/1] quit

```

On the MCE, create VLAN 30 and VLAN-interface 30, bind VLAN-interface 30 with VPN instance **vpn1** and configure an IPv6 address for the VLAN-interface 30.

```

[MCE] vlan 30
[MCE-vlan30] quit
[MCE] interface vlan-interface 30
[MCE-Vlan-interface30] ip binding vpn-instance vpn1
[MCE-Vlan-interface30] ipv6 address 30::1 64
[MCE-Vlan-interface30] quit

```

On the MCE, create VLAN 40 and VLAN-interface 40, bind VLAN-interface 40 with VPN instance **vpn2** and configure an IPv6 address for the VLAN-interface 40.

```

[MCE] vlan 40
[MCE-vlan40] quit
[MCE] interface vlan-interface 40
[MCE-Vlan-interface40] ip binding vpn-instance vpn2

```

```
[MCE-Vlan-interface40] ipv6 address 40::1 64
[MCE-Vlan-interface40] quit
```

On PE 1, create VLAN 30 and VLAN-interface 30, bind VLAN-interface 30 with VPN instance **vpn1** and configure an IPv6 address for the VLAN-interface 30.

```
[PE1] vlan 30
[PE1-vlan30] quit
[PE1] interface vlan-interface 30
[PE1-Vlan-interface30] ip binding vpn-instance vpn1
[PE1-Vlan-interface30] ipv6 address 30::2 64
[PE1-Vlan-interface30] quit
```

On PE 1, create VLAN 40 and VLAN-interface 40, bind VLAN-interface 40 with VPN instance **vpn2** and configure an IPv6 address for the VLAN-interface 40.

```
[PE1] vlan 40
[PE1-vlan40] quit
[PE1] interface vlan-interface 40
[PE1-Vlan-interface40] ip binding vpn-instance vpn2
[PE1-Vlan-interface40] ipv6 address 40::2 64
[PE1-Vlan-interface40] quit
```

Configure the IP address of the interface Loopback0 as 101.101.10.1 for the MCE and as 100.100.10.1 for PE 1. Specify the loopback interface address as the router ID for the MCE and PE 1. (Details not shown.)

Enable OSPFv3 process 10 on the MCE, bind the process to VPN instance **vpn1**, and redistribute the IPv6 static route of VPN 1.

```
[MCE] ospfv3 10 vpn-instance vpn1
[MCE-ospf-10] router-id 101.101.10.1
[MCE-ospf-10] import-route static
[MCE-ospf-10] quit
```

Enable OSPFv3 on VLAN-interface 30.

```
[MCE] interface vlan-interface 30
[MCE-Vlan-interface30] ospfv3 10 area 0.0.0.0
[MCE-Vlan-interface30] quit
```

On PE 1, enable OSPFv3 process 10 and bind the process to VPN instance **vpn1**.

```
[PE1] ospfv3 10 vpn-instance vpn1
[PE1-ospf-10] router-id 100.100.10.1
[PE1-ospf-10] quit
```

Enable OSPFv3 on VLAN-interface 30.

```
[PE1] interface vlan-interface 30
[PE1-Vlan-interface30] ospfv3 10 area 0.0.0.0
[PE1-Vlan-interface30] quit
```

On PE 1, display the routing table of VPN 1.

```
[PE1] display ipv6 routing-table vpn-instance vpn1
Routing Table : vpn1
                Destinations : 5                Routes : 5
```

```
Destination: ::1/128
```

```
Protocol : Direct
```

```

NextHop      : ::1                      Preference: 0
Interface    : InLoop0                  Cost       : 0

Destination: 30::/64                    Protocol   : Direct
NextHop      : 30::2                      Preference: 0
Interface    : Vlan30                    Cost       : 0

Destination: 30::2/128                   Protocol   : Direct
NextHop      : ::1                        Preference: 0
Interface    : InLoop0                    Cost       : 0

Destination: 2012:1::/64                  Protocol   : OSPFv3
NextHop      : FE80::202:FF:FE02:2        Preference: 150
Interface    : Vlan30                      Cost       : 1

Destination: FE80::/10                    Protocol   : Direct
NextHop      : ::                          Preference: 0
Interface    : NULL0                       Cost       : 0

```

The output shows that PE 1 has learned the private route of VPN 1 through OSPFv3.

Take similar procedures to configure OSPFv3 process 20 between the MCE and PE 1 and redistribute VPN 2's routes from RIPng process 20 into the OSPFv3 routing table of the MCE. The following output shows that PE 1 has learned the private route of VPN 2 through OSPFv3.

```

[PE1] display ipv6 routing-table vpn-instance vpn2
Routing Table : vpn2
                Destinations : 5          Routes : 5

```

```

Destination: ::1/128                      Protocol   : Direct
NextHop      : ::1                          Preference: 0
Interface    : InLoop0                       Cost       : 0

Destination: 40::/64                       Protocol   : Direct
NextHop      : 40::2                          Preference: 0
Interface    : Vlan40                         Cost       : 0

Destination: 40::2/128                     Protocol   : Direct
NextHop      : ::1                            Preference: 0
Interface    : InLoop0                         Cost       : 0

Destination: 2012::/64                      Protocol   : OSPFv3
NextHop      : FE80::200:FF:FE0F:5           Preference: 150
Interface    : Vlan40                          Cost       : 1

Destination: FE80::/10                      Protocol   : Direct
NextHop      : ::                            Preference: 0
Interface    : NULL0                           Cost       : 0

```

Now, the routing information of the two VPNs has been added into the routing tables on PE 1.

Index

A B C D E H I L M O P R S T

A

Applying IPsec policies for OSPFv3, [291](#)

Applying IPsec policies for RIPng, [268](#)

B

BGP configuration examples, [230](#)

BGP configuration task list, [201](#)

BGP overview, [187](#)

Binding an IS-IS process with MIBs, [161](#)

C

Configuring a large scale BGP network, [223](#)

Configuring a large-scale IPv6 BGP network, [342](#)

Configuring a routing policy, [365](#)

Configuring a static route, [7](#)

Configuring an IPv6 MCE, [418](#)

Configuring an IPv6 static route, [257](#)

Configuring an MCE, [396](#)

Configuring BFD for BGP, [228](#)

Configuring BFD for IPv6 BGP, [345](#)

Configuring BFD for IPv6 IS-IS, [315](#)

Configuring BFD for IS-IS, [161](#)

Configuring BFD for OSPF, [97](#)

Configuring BFD for OSPFv3, [291](#)

Configuring BFD for RIP, [36](#)

Configuring BFD for static routes, [8](#)

Configuring BGP basic functions, [202](#)

Configuring BGP GR, [227](#)

Configuring BGP route attributes, [211](#)

Configuring IPv6 BGP basic functions, [327](#)

Configuring IPv6 BGP route attributes, [334](#)

Configuring IPv6 IS-IS basic functions, [313](#)

Configuring IPv6 IS-IS routing information control, [314](#)

Configuring IS-IS basic functions, [143](#)

Configuring IS-IS FRR, [159](#)

Configuring IS-IS GR, [158](#)

Configuring IS-IS NSR, [159](#)

Configuring IS-IS routing information control, [145](#)

Configuring OSPF areas, [75](#)

Configuring OSPF FRR, [93](#)

Configuring OSPF Graceful Restart, [95](#)

Configuring OSPF network types, [77](#)

Configuring OSPF route control, [79](#)

Configuring OSPFv3 area parameters, [282](#)

Configuring OSPFv3 GR, [290](#)

Configuring OSPFv3 network types, [283](#)

Configuring OSPFv3 routing information control, [284](#)

Configuring PBR (using a PBR policy), [379](#)

Configuring PBR (using a QoS policy), [381](#)

Configuring RIP basic functions, [26](#)

Configuring RIP FRR, [36](#)

Configuring RIP route control, [28](#)

Configuring RIPng basic functions, [264](#)

Configuring RIPng route control, [264](#)

Configuring routing on an IPv6 MCE, [420](#)

Configuring routing on an MCE, [394](#)

Configuring routing on an MCE, [398](#)

Configuring static route FRR, [10](#)

Controlling route distribution and reception, [330](#)

Controlling route distribution and reception, [206](#)

Controlling route generation, [204](#)

D

Default IPv6 route, [257](#)

- Defining filters, [363](#)
- Displaying and maintaining a routing table, [4](#)
- Displaying and maintaining BGP, [228](#)
- Displaying and maintaining IPv6 BGP, [345](#)
- Displaying and maintaining IPv6 IS-IS, [315](#)
- Displaying and maintaining IPv6 static routes, [258](#)
- Displaying and maintaining IS-IS, [161](#)
- Displaying and maintaining MCE, [408](#)
- Displaying and maintaining OSPF, [97](#)
- Displaying and maintaining OSPFv3, [293](#)
- Displaying and maintaining PBR configuration, [383](#)
- Displaying and maintaining RIP, [37](#)
- Displaying and maintaining RIPng, [269](#)
- Displaying and maintaining static routes, [11](#)
- Displaying and maintaining the routing policy, [369](#)
- Displaying information about IPv6 MCE, [427](#)
- Dynamic routing protocols, [2](#)

E

- Enabling IS-IS SNMP trap, [160](#)
- Enabling logging of peer state changes, [228](#)
- Enabling OSPF, [73](#)
- Enabling OSPFv3, [281](#)
- Enabling trap, [227](#)
- Enhancing IS-IS network security, [157](#)

H

- Hardware compatibility, [377](#)
- Hardware compatibility, [6](#)
- Hardware compatibility, [187](#)
- Hardware compatibility, [1](#)
- Hardware compatibility, [21](#)
- Hardware compatibility, [361](#)
- Hardware compatibility, [278](#)
- Hardware compatibility, [313](#)
- Hardware compatibility, [325](#)
- Hardware compatibility, [257](#)
- Hardware compatibility, [261](#)
- Hardware compatibility, [129](#)
- Hardware compatibility, [55](#)

I

- Introduction, [6](#)
- Introduction to IPv6 IS-IS, [313](#)
- Introduction to OSPF, [55](#)
- Introduction to OSPFv3, [278](#)
- Introduction to PBR, [377](#)
- Introduction to RIPng, [261](#)
- Introduction to routing policy, [361](#)
- IPv6 BGP configuration examples, [347](#)
- IPv6 BGP configuration task list, [325](#)
- IPv6 BGP overview, [325](#)
- IPv6 IS-IS configuration examples, [316](#)
- IPv6 MCE configuration examples, [428](#)
- IPv6 static routes features, [257](#)
- IPv6 static routing configuration example, [258](#)
- IS-IS configuration examples, [162](#)
- IS-IS configuration task list, [142](#)
- IS-IS overview, [129](#)

L

- Load sharing, [3](#)

M

- MCE configuration examples, [409](#)
- MCE overview, [390](#)

O

- OSPF configuration examples, [99](#)
- OSPF configuration task list, [72](#)
- OSPFv3 configuration examples, [294](#)
- OSPFv3 configuration task list, [280](#)
- Overview, [257](#)
- Overview, [21](#)
- Overview, [418](#)
- Overview, [1](#)

P

- PBR configuration examples, [384](#)

R

- Resetting BGP connections, [427](#)

Resetting BGP connections, [407](#)
RIP configuration examples, [38](#)
RIP configuration task list, [25](#)
RIPng configuration examples, [270](#)
RIPng configuration task list, [263](#)
Route backup, [4](#)
Route recursion, [4](#)
Route redistribution, [4](#)
Routing policy configuration examples, [370](#)
Routing preference, [3](#)
Routing table, [1](#)

S

Static route configuration examples, [11](#)

T

Troubleshooting BGP, [256](#)
Troubleshooting IPv6 BGP configuration, [360](#)
Troubleshooting OSPF configuration, [127](#)
Troubleshooting OSPFv3 configuration, [311](#)
Troubleshooting RIP, [54](#)
Troubleshooting routing policy configuration, [376](#)
Tuning and optimizing BGP networks, [217](#)
Tuning and optimizing IPv6 BGP networks, [336](#)
Tuning and optimizing IS-IS networks, [150](#)
Tuning and optimizing OSPF networks, [85](#)
Tuning and optimizing OSPFv3 networks, [287](#)
Tuning and optimizing RIP networks, [32](#)
Tuning and optimizing the RIPng network, [266](#)

Contents

Multicast overview	1
Introduction to multicast	1
Information transmission techniques	1
Multicast features	3
Common notations in multicast	4
Multicast advantages and applications	4
Multicast models	5
Multicast architecture	5
Multicast addresses	6
Multicast protocols	9
Multicast packet forwarding mechanism	11
Multicast support for VPNs	11
Introduction to VPN instances	11
Multicast application in VPNs	12
Configuring IGMP snooping	13
Overview	13
Basic concepts in IGMP snooping	13
How IGMP snooping works	15
IGMP snooping proxying	16
Protocols and standards	18
IGMP snooping configuration task list	18
Configuring basic IGMP snooping functions	19
Enabling IGMP snooping	19
Specifying the version of IGMP snooping	19
Configuring static multicast MAC address entries	20
Configuring IGMP snooping port functions	21
Configuration prerequisites	21
Setting aging timers for dynamic ports	21
Configuring static ports	22
Configuring a port as a simulated member host	23
Enabling fast-leave processing	23
Disabling a port from becoming a dynamic router port	24
Configuring IGMP snooping querier	25
Configuration prerequisites	25
Enabling IGMP snooping querier	25
Configuring parameters for IGMP queries and responses	26
Configuring the source IP addresses for IGMP queries	27
Configuring IGMP snooping proxying	27
Configuration prerequisites	27
Enabling IGMP snooping proxying	27
Configuring a source IP address for the IGMP messages sent by the proxy	28
Configuring an IGMP snooping policy	28
Configuration prerequisites	28
Configuring a multicast group filter	28
Configuring multicast source port filtering	29
Enabling dropping unknown multicast data	30
Configuring IGMP report suppression	31
Setting the maximum number of multicast groups that a port can join	31

Enabling multicast group replacement.....	32
Setting the 802.1p precedence for IGMP messages	33
Configuring a multicast user control policy	33
Enabling the IGMP snooping host tracking function	34
Setting the DSCP value for IGMP messages.....	35
Displaying and maintaining IGMP snooping.....	35
IGMP snooping configuration examples	36
Group policy and simulated joining configuration example.....	36
Static port configuration example	38
IGMP snooping querier configuration example.....	41
IGMP snooping proxying configuration example.....	43
Multicast source and user control policy configuration example.....	46
Troubleshooting IGMP snooping	51
Layer 2 multicast forwarding cannot function	51
Configured multicast group policy fails to take effect	51
Appendix	51
Processing of multicast protocol messages	51
Configuring PIM snooping	53
Overview.....	53
Configuring PIM snooping	54
Displaying and maintaining PIM snooping	55
PIM snooping configuration example.....	55
Troubleshooting PIM snooping	58
PIM snooping does not work	58
Some downstream PIM-capable routers cannot receive multicast data	58
Configuring multicast VLANs.....	59
Overview.....	59
Multicast VLAN configuration task list	61
Configuring a sub-VLAN-based multicast VLAN.....	61
Configuration prerequisites	61
Configuration guidelines	61
Configuration procedure	61
Configuring a port-based multicast VLAN.....	62
Configuration prerequisites	62
Configuring user port attributes	62
Configuring multicast VLAN ports	63
Displaying and maintaining multicast VLAN.....	64
Multicast VLAN configuration examples.....	64
Sub-VLAN-based multicast VLAN configuration example	64
Port-based multicast VLAN configuration example	68
Configuring multicast routing and forwarding (available only on the HP 5500 EI).....	72
Overview.....	72
RPF check mechanism	72
Static multicast routes.....	74
Multicast traceroute.....	76
Configuration task list.....	76
Enabling IP multicast routing.....	77
Configuring multicast routing and forwarding.....	77
Configuration prerequisites	77
Configuring static multicast routes	78
Configuring a multicast routing policy	78
Configuring a multicast forwarding range	79
Configuring the multicast forwarding table size	79

Tracing a multicast path	80
Displaying and maintaining multicast routing and forwarding	81
Configuration examples	82
Changing an RPF route	82
Creating an RPF route	84
Troubleshooting multicast routing and forwarding	86
Static multicast route failure	86
Multicast data fails to reach receivers	86
Configuring IGMP (available only on the HP 5500 EI)	88
Overview	88
IGMP versions	88
Introduction to IGMPv1	88
Enhancements in IGMPv2	90
Enhancements in IGMPv3	90
IGMP SSM mapping	92
IGMP proxying	93
IGMP support for VPNs	94
Protocols and standards	94
IGMP configuration task list	94
Configuring basic IGMP functions	95
Configuration prerequisites	95
Enabling IGMP	95
Configuring IGMP versions	96
Configuring static joining	97
Configuring a multicast group filter	97
Setting the maximum number of multicast groups that an interface can join	98
Adjusting IGMP performance	98
Configuration prerequisites	98
Configuring IGMP message options	98
Configuring IGMP query and response parameters	99
Configuring IGMP fast-leave processing	101
Enabling the IGMP host tracking function	102
Setting the DSCP value for IGMP messages	103
Configuring IGMP SSM mapping	103
Enabling SSM mapping	103
Configuring SSM mappings	104
Configuring IGMP proxying	104
Configuration prerequisites	104
Enabling IGMP proxying	104
Configuring multicast forwarding on a downstream interface	105
Displaying and maintaining IGMP	105
IGMP configuration examples	108
Basic IGMP functions configuration example	108
SSM mapping configuration example	110
IGMP proxying configuration example	113
Troubleshooting IGMP	114
No membership information on the receiver-side router	114
Inconsistent memberships on routers on the same subnet	115
Configuring PIM (available only on the HP 5500 EI)	116
PIM overview	116
PIM-DM overview	116
PIM-SM overview	119
BIDIR-PIM overview	125

Administrative scoping overview	128
PIM-SSM overview	130
Relationships among PIM protocols	131
PIM support for VPNs	132
Protocols and standards	132
Configuring PIM-DM	132
PIM-DM configuration task list	132
Configuration prerequisites	133
Enabling PIM-DM	133
Enabling state-refresh capability	134
Configuring state-refresh parameters	134
Configuring PIM-DM graft retry period	135
Configuring PIM-SM	135
PIM-SM configuration task list	135
Configuration prerequisites	136
Enabling PIM-SM	137
Configuring an RP	137
Configuring a BSR	140
Configuring administrative scoping	143
Configuring multicast source registration	145
Disabling SPT switchover	146
Configuring BIDIR-PIM	147
BIDIR-PIM configuration task list	147
Configuration prerequisites	147
Enabling PIM-SM	148
Enabling BIDIR-PIM	149
Configuring an RP	149
Configuring a BSR	151
Configuring administrative scoping	153
Configuring PIM-SSM	154
PIM-SSM configuration task list	155
Configuration prerequisites	155
Enabling PIM-SM	155
Configuring the SSM group range	156
Configuring PIM common features	156
PIM common feature configuration task list	157
Configuration prerequisites	157
Configuring a multicast data filter	158
Configuring a hello message filter	158
Configuring PIM hello options	158
Configuring the prune delay	160
Configuring PIM common timers	160
Configuring join/prune message sizes	162
Configuring PIM to work with BFD	162
Setting the DSCP value for PIM messages	163
Displaying and maintaining PIM	163
PIM configuration examples	165
PIM-DM configuration example	165
PIM-SM non-scoped zone configuration example	168
PIM-SM admin-scope zone configuration example	173
BIDIR-PIM configuration example	179
PIM-SSM configuration example	183
Troubleshooting PIM	186
A multicast distribution tree cannot be built correctly	186

Multicast data abnormally terminated on an intermediate router	187
RPs cannot join SPT in PIM-SM	187
RPT establishment failure or source registration failure in PIM-SM	188
Configuring MSDP (available only on the HP 5500 EI)	189
Overview	189
How MSDP works	189
MSDP support for VPNs	194
Protocols and standards	194
MSDP configuration task list	194
Configuring basic MSDP functions	195
Configuration prerequisites	195
Enabling MSDP	195
Creating an MSDP peer connection	196
Configuring a static RPF peer	196
Configuring an MSDP peer connection	197
Configuration prerequisites	197
Configuring MSDP peer description	197
Configuring an MSDP mesh group	197
Configuring MSDP peer connection control	198
Configuring SA messages related parameters	199
Configuration prerequisites	199
Configuring SA message content	199
Configuring SA request messages	200
Configuring SA message filtering rules	200
Configuring the SA cache mechanism	201
Displaying and maintaining MSDP	202
MSDP configuration examples	203
PIM-SM Inter-domain multicast configuration	203
Inter-AS multicast configuration by leveraging static RPF peers	207
Anycast RP configuration	211
SA message filtering configuration	215
Troubleshooting MSDP	219
MSDP peers stay in down state	219
No SA entries in the switch's SA cache	219
Inter-RP communication faults in Anycast RP application	220
Configuring MBGP (available only on the HP 5500 EI)	221
MBGP overview	221
Protocols and standards	221
MBGP configuration task list	221
Configuring basic MBGP functions	222
Configuration prerequisites	222
Configuration procedure	222
Controlling route advertisement and reception	223
Configuration prerequisites	223
Configuring MBGP route redistribution	223
Configuring default route redistribution into MBGP	223
Configuring MBGP route summarization	224
Advertising a default route to an IPv4 MBGP peer or peer group	224
Configuring outbound MBGP route filtering	225
Configuring inbound MBGP route filtering	226
Configuring MBGP route dampening	227
Configuring MBGP route attributes	228
Configuration prerequisites	228

Configuring MBGP route preferences	228
Configuring the default local preference	228
Configuring the MED attribute	229
Configuring the next hop attribute	229
Configuring the AS-PATH attributes	230
Tuning and optimizing MBGP networks	230
Configuration prerequisites	230
Configuring MBGP soft reset	230
Enabling the MBGP ORF capability	231
Configuring the maximum number of MBGP routes for load balancing	232
Configuring a large scale MBGP network	233
Configuration prerequisites	233
Configuring IPv4 MBGP peer groups	233
Configuring MBGP community	234
Configuring an MBGP route reflector	234
Displaying and maintaining MBGP	235
Displaying MBGP	235
Resetting MBGP connections	236
Clearing MBGP information	236
MBGP configuration example	237
Configuring MLD snooping	241
Overview	241
Basic concepts in MLD snooping	241
How MLD snooping works	243
MLD snooping proxying	244
Protocols and standards	246
MLD snooping configuration task list	246
Configuring basic MLD snooping functions	247
Configuration prerequisites	247
Enabling MLD snooping	247
Specifying the version of MLD snooping	248
Configuring IPv6 static multicast MAC address entries	248
Configuring MLD snooping port functions	249
Configuration prerequisites	249
Configuring aging timers for dynamic ports	249
Configuring static ports	250
Configuring a port as a simulated member host	251
Enabling fast-leave processing	251
Disabling a port from becoming a dynamic router port	252
Configuring MLD snooping querier	253
Configuration prerequisites	253
Enabling MLD snooping querier	253
Configuring parameters for MLD queries and responses	254
Configuring the source IPv6 addresses for MLD queries	255
Configuring MLD snooping proxying	255
Configuration prerequisites	255
Enabling MLD snooping proxying	255
Configuring the source IPv6 addresses for the MLD messages sent by the proxy	256
Configuring an MLD snooping policy	256
Configuration prerequisites	256
Configuring an IPv6 multicast group filter	256
Configuring IPv6 multicast source port filtering	257
Enabling dropping unknown IPv6 multicast data	258
Configuring MLD report suppression	259

Setting the maximum number of multicast groups that a port can join	259
Enabling IPv6 multicast group replacement	260
Setting the 802.1p precedence for MLD messages	261
Configuring an IPv6 multicast user control policy	261
Enabling the MLD snooping host tracking function	262
Setting the DSCP value for MLD messages	263
Displaying and maintaining MLD snooping	263
MLD snooping configuration examples	264
IPv6 group policy and simulated joining configuration example	264
Static port configuration example	266
MLD snooping querier configuration example	270
MLD snooping proxying configuration example	271
IPv6 multicast source and user control policy configuration example	274
Troubleshooting MLD snooping	279
Layer 2 multicast forwarding cannot function	279
Configured IPv6 multicast group policy fails to take effect	279
Appendix	279
Processing of IPv6 multicast protocol messages	279
Configuring IPv6 PIM snooping	281
Overview	281
Configuring IPv6 PIM snooping	282
Displaying and maintaining IPv6 PIM snooping	283
IPv6 PIM snooping configuration example	283
Troubleshooting IPv6 PIM snooping	286
IPv6 PIM snooping does not work	286
Some downstream IPv6 PIM-capable routers cannot receive multicast data	286
Configuring IPv6 multicast VLANs	287
Overview	287
IPv6 multicast VLAN configuration task list	289
Configuring a sub-VLAN-based IPv6 multicast VLAN	289
Configuration prerequisites	289
Configuration guidelines	289
Configuration procedure	289
Configuring a port-based IPv6 multicast VLAN	290
Configuration prerequisites	290
Configuring user port attributes	290
Configuring IPv6 multicast VLAN ports	291
Displaying and maintaining IPv6 multicast VLAN	292
IPv6 multicast VLAN configuration examples	292
Sub-VLAN-based multicast VLAN configuration example	292
Port-based multicast VLAN configuration example	296
Configuring IPv6 multicast routing and forwarding (available only on the HP 5500 EI)	300
Overview	300
RPF check mechanism	300
Configuration task list	302
Enabling IPv6 multicast routing	303
Configuring IPv6 multicast routing and forwarding	303
Configuration prerequisites	303
Configuring an IPv6 multicast routing policy	303
Configuring an IPv6 multicast forwarding range	303
Configuring the IPv6 multicast forwarding table size	304
Displaying and maintaining IPv6 multicast routing and forwarding	305
Troubleshooting IPv6 multicast policy configuration	306

Abnormal termination of IPv6 multicast data	306
Configuring MLD (available only on the HP 5500 EI)	308
Overview	308
MLD versions	308
How MLDv1 works	308
How MLDv2 works	310
MLD messages	311
MLD SSM mapping	314
MLD proxying	315
Protocols and standards	315
MLD configuration task list	316
Configuring basic MLD functions	316
Configuration prerequisites	316
Enabling MLD	317
Configuring the MLD version	317
Configuring static joining	318
Configuring an IPv6 multicast group filter	318
Setting the maximum number of IPv6 multicast groups that an interface can join	319
Adjusting MLD performance	319
Configuration prerequisites	319
Configuring MLD message options	319
Configuring MLD query and response parameters	320
Configuring MLD fast-leave processing	323
Enabling the MLD host tracking function	323
Setting the DSCP value for MLD messages	324
Configuring MLD SSM mapping	324
Configuration prerequisites	324
Enabling MLD SSM mapping	324
Configuring MLD SSM mappings	325
Configuring MLD proxying	325
Configuration prerequisites	325
Enabling MLD proxying	325
Configuring IPv6 multicast forwarding on a downstream interface	326
Displaying and maintaining MLD	326
MLD configuration examples	328
Basic MLD functions configuration example	328
MLD SSM mapping configuration example	330
MLD proxying configuration example	333
Troubleshooting MLD	335
No member information on the receiver-side router	335
Inconsistent memberships on routers on the same subnet	336
Configuring IPv6 PIM (available only on the HP 5500 EI)	337
Overview	337
IPv6 PIM-DM overview	337
IPv6 PIM-SM overview	340
IPv6 BIDIR-PIM overview	346
IPv6 administrative scoping overview	349
IPv6 PIM-SSM overview	351
Relationships among IPv6 PIM protocols	353
Protocols and standards	353
Configuring IPv6 PIM-DM	353
IPv6 PIM-DM configuration task list	353
Configuration prerequisites	354

Enabling IPv6 PIM-DM	354
Enabling state-refresh capability	354
Configuring state refresh parameters	355
Configuring IPv6 PIM-DM graft retry period	355
Configuring IPv6 PIM-SM	356
IPv6 PIM-SM configuration task list	356
Configuration prerequisites	356
Enabling IPv6 PIM-SM	357
Configuring an RP	357
Configuring a BSR	360
Configuring IPv6 administrative scoping	363
Configuring IPv6 multicast source registration	365
Disabling SPT switchover	366
Configuring IPv6 BIDIR-PIM	366
IPv6 BIDIR-PIM configuration task list	366
Configuration prerequisites	367
Enabling IPv6 PIM-SM	367
Enabling IPv6 BIDIR-PIM	368
Configuring an RP	368
Configuring a BSR	370
Configuring IPv6 administrative scoping	372
Configuring IPv6 PIM-SSM	373
IPv6 PIM-SSM configuration task list	373
Configuration prerequisites	373
Enabling IPv6 PIM-SM	374
Configuring the IPv6 SSM group range	374
Configuring IPv6 PIM common features	375
IPv6 PIM common feature configuration task list	375
Configuration prerequisites	375
Configuring an IPv6 multicast data filter	376
Configuring a hello message filter	376
Configuring IPv6 PIM hello options	377
Configuring the prune delay	378
Configuring IPv6 PIM common timers	379
Configuring join/prune message sizes	380
Configuring IPv6 PIM to work with BFD	381
Setting the DSCP value for IPv6 PIM messages	381
Displaying and maintaining IPv6 PIM	381
IPv6 PIM configuration examples	383
IPv6 PIM-DM configuration example	383
IPv6 PIM-SM non-scoped zone configuration example	386
IPv6 PIM-SM admin-scope zone configuration example	391
IPv6 BIDIR-PIM configuration example	403
IPv6 PIM-SSM configuration example	408
Troubleshooting IPv6 PIM configuration	410
Failure to build a multicast distribution tree correctly	410
IPv6 multicast data abnormally terminated on an intermediate router	411
RPS cannot join SPT in IPv6 PIM-SM	412
RPT establishment failure or source registration failure in IPv6 PIM-SM	412
Configuring IPv6 MBGP (available only on the HP 5500 EI)	414
IPv6 MBGP overview	414
Configuring basic IPv6 MBGP functions	415
Configuration prerequisites	415
Configuring an IPv6 MBGP peer	415

Configuring a preferred value for routes from a peer or a peer group	415
Controlling route distribution and reception	416
Configuration prerequisites	416
Injecting a local IPv6 MBGP route	416
Configuring IPv6 MBGP route redistribution	416
Configuring IPv6 MBGP route summarization	417
Advertising a default route to a peer or peer group	417
Configuring outbound IPv6 MBGP route filtering	418
Configuring inbound IPv6 MBGP route filtering	418
Configuring IPv6 MBGP route dampening	419
Configuring IPv6 MBGP route attributes	419
Configuration prerequisites	420
Configuring IPv6 MBGP route preferences	420
Configuring the default local preference	420
Configuring the MED attribute	420
Configuring the next hop attribute	421
Configuring the as_path attribute	421
Tuning and optimizing IPv6 MBGP networks	422
Configuration prerequisites	422
Configuring IPv6 MBGP soft reset	422
Enabling the IPv6 MBGP orf capability	423
Configuring the maximum number of equal-cost routes for load-balancing	424
Configuring a large scale IPv6 MBGP network	425
Configuration prerequisites	425
Configuring an IPv6 MBGP peer group	425
Configuring IPv6 MBGP community	425
Configuring an IPv6 MBGP route reflector	426
Displaying and maintaining IPv6 MBGP	427
Displaying IPv6 MBGP	427
Resetting IPv6 MBGP connections	428
Clearing IPv6 MBGP information	428
IPv6 MBGP configuration example	429

Index	432
-------------	-----

Multicast overview

Introduction to multicast

As a technique that coexists with unicast and broadcast, the multicast technique effectively addresses the issue of point-to-multipoint data transmission. By enabling high-efficiency point-to-multipoint data transmission over a network, multicast greatly saves network bandwidth and reduces network load.

By using multicast technology, a network operator can easily provide new value-added services, such as live webcasting, web TV, distance learning, telemedicine, web radio, real time video conferencing, and other bandwidth-critical and time-critical information services.

The term "router " in this document refers to both routers and Layer 3 switches.

Unless otherwise stated, the term "multicast" in this document refers to IP multicast.

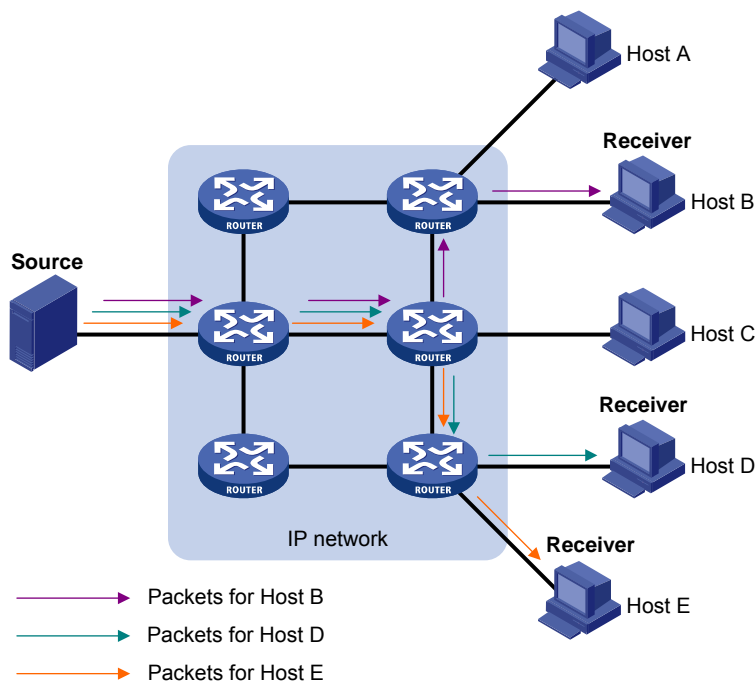
Information transmission techniques

The information transmission techniques include unicast, broadcast, and multicast.

Unicast

In unicast transmission, the information source must send a separate copy of information to each host that needs the information.

Figure 1 Unicast transmission



In [Figure 1](#), assume that Host B, Host D and Host E need the information. A separate transmission channel must be established from the information source to each of these hosts.

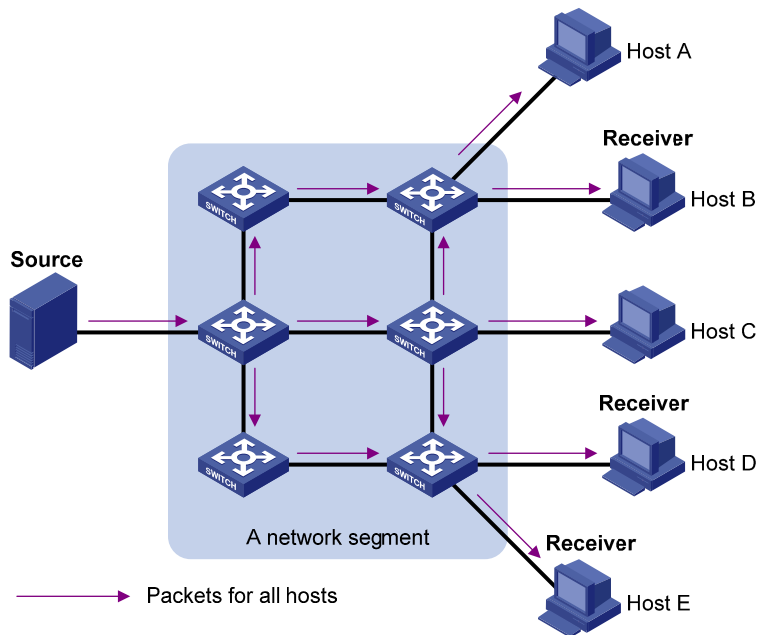
In unicast transmission, the traffic transmitted over the network is proportional to the number of hosts that need the information. If a large number of hosts need the information, the information source must send a separate copy of the same information to each of these hosts. Sending many copies can place a tremendous pressure on the information source and the network bandwidth.

Unicast is not suitable for batch transmission of information.

Broadcast

In broadcast transmission, the information source sends information to all hosts on the subnet, even if some hosts do not need the information.

Figure 2 Broadcast transmission



In Figure 2, assume that only Host B, Host D, and Host E need the information. If the information is broadcast to the subnet, Host A and Host C also receive it. In addition to information security issues, broadcasting to hosts that do not need the information also causes traffic flooding on the same subnet.

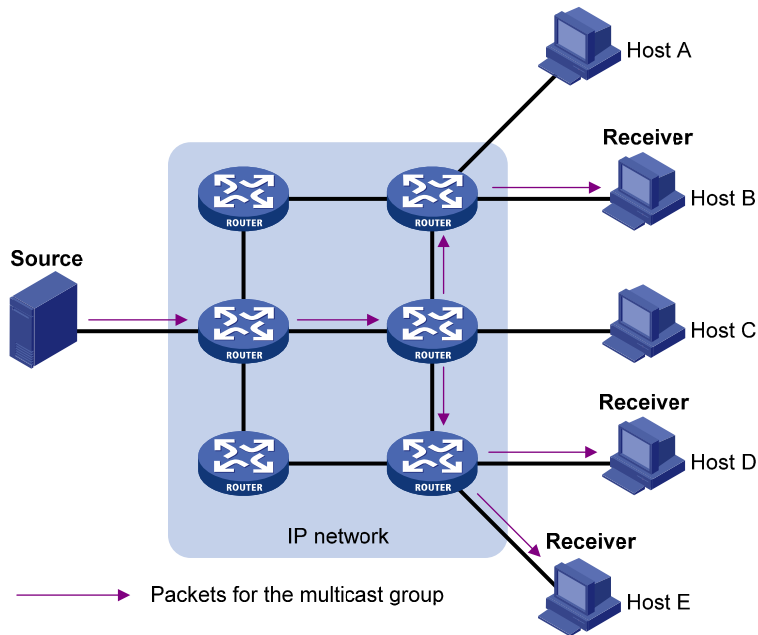
Broadcast is disadvantageous in transmitting data to specific hosts. Moreover, broadcast transmission is a significant waste of network resources.

Multicast

Unicast and broadcast techniques cannot provide point-to-multipoint data transmissions with the minimum network consumption.

Multicast transmission can solve this problem. When some hosts on the network need multicast information, the information sender, or multicast source, sends only one copy of the information. Multicast distribution trees are built through multicast routing protocols, and the packets are replicated only on nodes where the trees branch.

Figure 3 Multicast transmission



The multicast source sends only one copy of the information to a multicast group. Host B, Host D and Host E, which are receivers of the information, must join the multicast group. The routers on the network duplicate and forward the information based on the distribution of the group members. Finally, the information is correctly delivered to Host B, Host D, and Host E.

To summarize, multicast has the following advantages:

- **Advantages over unicast**—Because multicast traffic flows to the farthest-possible node from the source before it is replicated and distributed, an increase in the number of hosts does not increase the load of the source or remarkably add to the usage of network resources.
- **Advantages over broadcast**—Because multicast data is sent only to the receivers that need it, multicast uses network bandwidth reasonably and enhances network security. In addition, data broadcast is confined to the same subnet, but multicast is not.

Multicast features

- A multicast group is a multicast receiver set identified by an IP multicast address. Hosts join a multicast group to become members of the multicast group before they can receive the multicast data addressed to that multicast group. Typically, a multicast source does not need to join a multicast group.
- An information sender is called a "multicast source". A multicast source can send data to multiple multicast groups at the same time, and multiple multicast sources can send data to the same multicast group at the same time.
- All hosts that have joined a multicast group become members of the multicast group. The group memberships are dynamic. Hosts can join or leave multicast groups at any time. Multicast groups are not subject to geographic restrictions.
- Routers or Layer 3 switches that support Layer 3 multicast are called "multicast routers" or "Layer 3 multicast devices". In addition to providing the multicast routing function, a multicast router can also manage multicast group memberships on stub subnets with attached group members. A multicast router itself can be a multicast group member.

For a better understanding of the multicast concept, you can compare multicast transmission to the transmission of TV programs.

Table 1 Comparing TV transmission and multicast transmission

TV transmission	Multicast transmission
A TV station transmits a TV program through a channel.	A multicast source sends multicast data to a multicast group.
A user tunes the TV set to the channel.	A receiver joins the multicast group.
The user starts to watch the TV program transmitted by the TV station via the channel.	The receiver starts to receive the multicast data that the source is sending to the multicast group.
The user turns off the TV set or tunes to another channel.	The receiver leaves the multicast group or joins another group.

Common notations in multicast

The following notations are commonly used in multicast transmission:

- **(*, G)**—Indicates a rendezvous point tree (RPT), or a multicast packet that any multicast source sends to multicast group G. Here, the asterisk represents any multicast source, and "G" represents a specific multicast group.
- **(S, G)**—Indicates a shortest path tree (SPT), or a multicast packet that multicast source S sends to multicast group G. Here, "S" represents a specific multicast source, and "G" represents a specific multicast group.

For more information about the concepts RPT and SPT, see "[Configuring PIM \(available only on the HP 5500 EI\)](#)" and "[Configuring IPv6 PIM \(available only on the HP 5500 EI\)](#)."

Multicast advantages and applications

Multicast advantages

Advantages of the multicast technique include the following:

- **Enhanced efficiency**—Reduces the processor load of information source servers and network devices.
- **Optimal performance**—Reduces redundant traffic.
- **Distributed application**—Enables point-to-multipoint applications at the price of minimum network resources.

Multicast applications

The scenarios in which the multicast technique can be effectively applied are:

- Multimedia and streaming applications, such as web TV, web radio, and real time video/audio conferencing
- Communication for training and cooperative operations, such as distance learning and telemedicine
- Data warehouse and financial applications (stock quotes)
- Any other point-to-multipoint application for data distribution

Multicast models

Based on how the receivers treat the multicast sources, the multicast models include any-source multicast (ASM), source-filtered multicast (SFM), and source-specific multicast (SSM).

ASM model

In the ASM model, any sender can send information to a multicast group as a multicast source, and receivers can join a multicast group (identified by a group address) and obtain multicast information addressed to that multicast group. In this model, receivers do not know the positions of the multicast sources in advance. However, they can join or leave the multicast group at any time.

SFM model

The SFM model is derived from the ASM model. To a sender, the two models appear to have the same multicast membership architecture.

The SFM model functionally extends the ASM model. The upper-layer software checks the source address of received multicast packets and permits or denies multicast traffic from specific sources. Therefore, receivers can receive the multicast data from only part of the multicast sources. To a receiver, multicast sources are not all valid; they are filtered.

SSM model

Users might be interested in the multicast data from only certain multicast sources. The SSM model provides a transmission service that enables users to specify the multicast sources that they are interested in at the client side.

The main difference between the SSM model and the ASM model is that in the SSM model, receivers have already determined the locations of the multicast sources by some other means. In addition, the SSM model uses a multicast address range that is different from that of the ASM/SFM model, and dedicated multicast forwarding paths are established between receivers and the specified multicast sources.

Multicast architecture

IP multicast addresses the following questions:

- Where should the multicast source transmit information to? (multicast addressing)
- What receivers exist on the network? (host registration)
- Where is the multicast source that will provide data to the receivers? (multicast source discovery)
- How should information be transmitted to the receivers? (multicast routing)

IP multicast is an end-to-end service. The multicast architecture involves the following parts:

- **Addressing mechanism**—A multicast source sends information to a group of receivers through a multicast address.
- **Host registration**—Receiver hosts can join and leave multicast groups dynamically. This mechanism is the basis for management of group memberships.
- **Multicast routing**—A multicast distribution tree (namely, a forwarding path tree for multicast data on the network) is constructed for delivering multicast data from a multicast source to receivers.
- **Multicast applications**—A software system that supports multicast applications, such as video conferencing, must be installed on multicast sources and receiver hosts. The TCP/IP stack must support reception and transmission of multicast data.

Multicast addresses

Network-layer multicast addresses (multicast IP addresses) enables communication between multicast sources and multicast group members. In addition, a technique must be available to map multicast IP addresses to link-layer multicast MAC addresses.

IP multicast addresses

- IPv4 multicast addresses
Internet Assigned Numbers Authority (IANA) assigned the Class D address space (224.0.0.0 to 239.255.255.255) to IPv4 multicast.

Table 2 Class D IP address blocks and description

Address block	Description
224.0.0.0 to 224.0.0.255	Reserved permanent group addresses. The IP address 224.0.0.0 is reserved. Other IP addresses can be used by routing protocols and for topology searching, protocol maintenance, and so on. Table 3 lists common permanent group addresses. A packet destined for an address in this block will not be forwarded beyond the local subnet regardless of the Time to Live (TTL) value in the IP header.
224.0.1.0 to 238.255.255.255	Globally scoped group addresses. This block includes the following types of designated group addresses: <ul style="list-style-type: none">• 232.0.0.0/8—SSM group addresses, and• 233.0.0.0/8—Glop group addresses.
239.0.0.0 to 239.255.255.255	Administratively scoped multicast addresses. These addresses are considered locally unique rather than globally unique, and can be reused in domains administered by different organizations without causing conflicts. For more information, see RFC 2365.

NOTE:

"Glop" is a mechanism for assigning multicast addresses between different autonomous systems (ASs). By filling an AS number into the middle two bytes of 233.0.0.0, you get 255 multicast addresses for that AS. For more information, see RFC 2770.

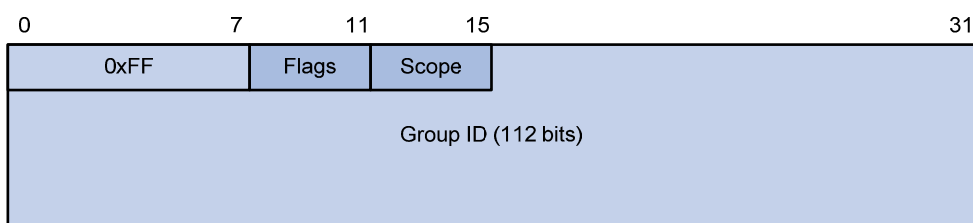
Table 3 Some reserved multicast addresses

Address	Description
224.0.0.1	All systems on this subnet, including hosts and routers
224.0.0.2	All multicast routers on this subnet
224.0.0.3	Unassigned
224.0.0.4	Distance Vector Multicast Routing Protocol (DVMRP) routers
224.0.0.5	Open Shortest Path First (OSPF) routers
224.0.0.6	OSPF designated routers and backup designated routers
224.0.0.7	Shared Tree (ST) routers
224.0.0.8	ST hosts
224.0.0.9	Routing Information Protocol version 2 (RIPv2) routers
224.0.0.11	Mobile agents

Address	Description
224.0.0.12	Dynamic Host Configuration Protocol (DHCP) server/relay agent
224.0.0.13	All Protocol Independent Multicast (PIM) routers
224.0.0.14	Resource Reservation Protocol (RSVP) encapsulation
224.0.0.15	All Core-Based Tree (CBT) routers
224.0.0.16	Designated Subnetwork Bandwidth Management (SBM)
224.0.0.17	All SBMs
224.0.0.18	Virtual Router Redundancy Protocol (VRRP)

- IPv6 multicast addresses

Figure 4 IPv6 multicast format



The following describes the fields of an IPv6 multicast address:

- **0xFF**—The most significant eight bits are 11111111, which indicates that this address is an IPv6 multicast address.
- **Flags**—The Flags field contains four bits.

Figure 5 Flags field format



Table 4 Flags field description

Bit	Description
0	Reserved, set to 0
R	<ul style="list-style-type: none"> • When set to 0, it indicates that this address is an IPv6 multicast address without an embedded RP address • When set to 1, it indicates that this address is an IPv6 multicast address with an embedded RP address (the P and T bits must also be set to 1)
P	<ul style="list-style-type: none"> • When set to 0, it indicates that this address is an IPv6 multicast address not based on a unicast prefix • When set to 1, it indicates that this address is an IPv6 multicast address based on a unicast prefix (the T bit must also be set to 1)
T	<ul style="list-style-type: none"> • When set to 0, it indicates that this address is an IPv6 multicast address permanently-assigned by IANA • When set to 1, it indicates that this address is a transient, or dynamically assigned IPv6 multicast address

- **Scope**—The Scope field contains four bits, which indicate the scope of the IPv6 internetwork for which the multicast traffic is intended.

Table 5 Values of the Scope field

Value	Meaning
0, F	Reserved
1	Interface-local scope
2	Link-local scope
3	Subnet-local scope
4	Admin-local scope
5	Site-local scope
6, 7, 9 through D	Unassigned
8	Organization-local scope
E	Global scope

- **Group ID**—The Group ID field contains 112 bits. It uniquely identifies an IPv6 multicast group in the scope that the Scope field defines.

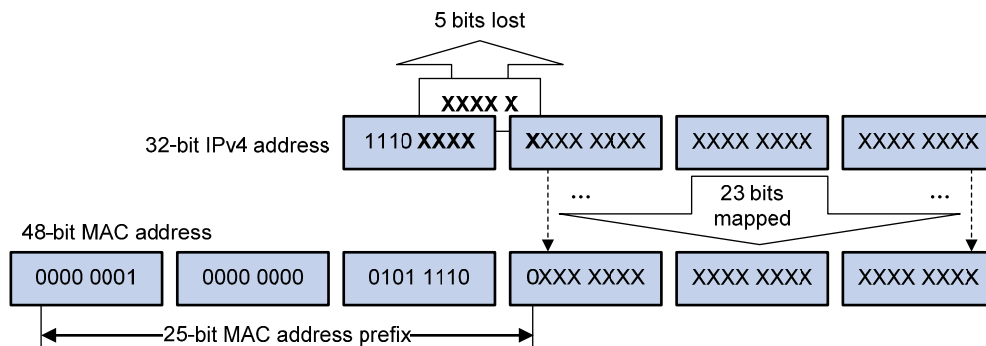
Ethernet multicast MAC addresses

A multicast MAC address identifies a group of receivers at the data link layer.

- IPv4 multicast MAC addresses

As defined by IANA, the most significant 24 bits of an IPv4 multicast MAC address are 0x01005E. Bit 25 is 0, and the other 23 bits are the least significant 23 bits of a multicast IPv4 address.

Figure 6 IPv4-to-MAC address mapping

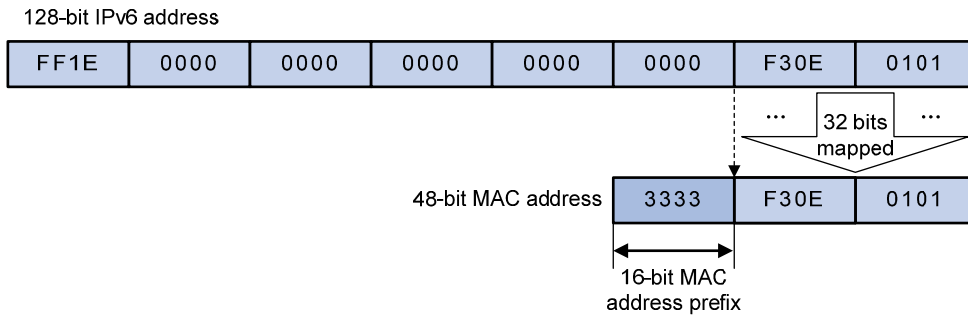


The most significant four bits of a multicast IPv4 address are 1110, which indicates that this address is a multicast address. Only 23 bits of the remaining 28 bits are mapped to a MAC address, so five bits of the multicast IPv4 address are lost. As a result, 32 multicast IPv4 addresses map to the same IPv4 multicast MAC address. Therefore, in Layer 2 multicast forwarding, a switch might receive some multicast data destined for other IPv4 multicast groups. The upper layer must filter such redundant data.

- IPv6 multicast MAC addresses

The most significant 16 bits of an IPv6 multicast MAC address are 0x3333. The least significant 32 bits are the least significant 32 bits of a multicast IPv6 address.

Figure 7 An example of IPv6-to-MAC address mapping



Multicast protocols

Generally, Layer 3 multicast refers to IP multicast working at the network layer. The corresponding multicast protocols are Layer 3 multicast protocols, which include IGMP, MLD, PIM, IPv6 PIM, MSDP, MBGP, and IPv6 MBGP. Layer 2 multicast refers to IP multicast working at the data link layer. The corresponding multicast protocols are Layer 2 multicast protocols, which include IGMP snooping, MLD snooping, PIM snooping, IPv6 PIM snooping, multicast VLAN, and IPv6 multicast VLAN.

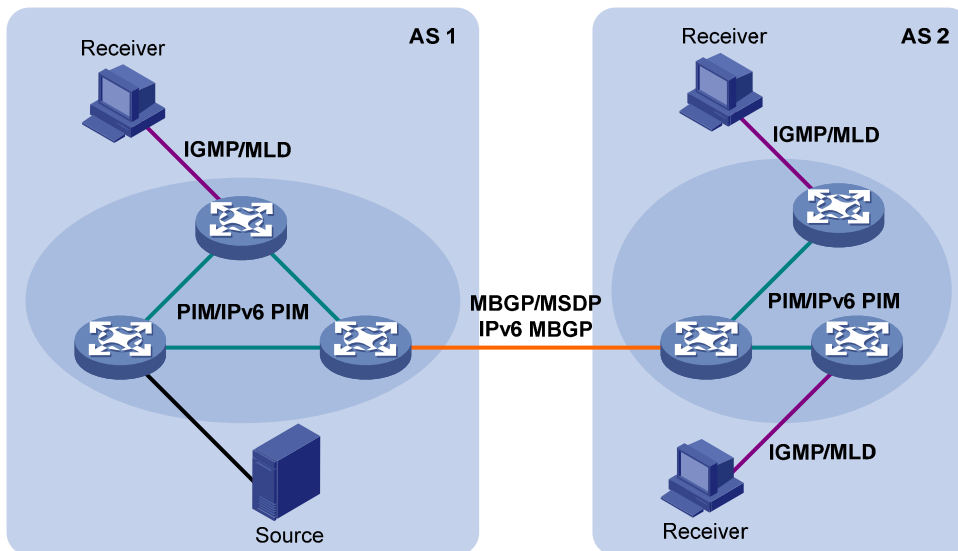
IGMP snooping, PIM snooping, multicast VLAN, IGMP, PIM, MSDP, and MBGP are for IPv4, and MLD snooping, IPv6 PIM snooping, IPv6 multicast VLAN, MLD, IPv6 PIM, and IPv6 MBGP are for IPv6.

This section provides only general descriptions about applications and functions of the Layer 2 and Layer 3 multicast protocols in a network. For more information about these protocols, see the related chapters.

Layer 3 multicast protocols

Layer 3 multicast protocols include multicast group management protocols and multicast routing protocols.

Figure 8 Positions of Layer 3 multicast protocols



- Multicast group management protocols
Typically, the Internet Group Management Protocol (IGMP) or Multicast Listener Discovery Protocol (MLD) is used between hosts and Layer 3 multicast devices that directly connect to the hosts. These

protocols define the mechanism of establishing and maintaining group memberships between hosts and Layer 3 multicast devices.

- Multicast routing protocols

A multicast routing protocol runs on Layer 3 multicast devices to establish and maintain multicast routes and forward multicast packets correctly and efficiently. Multicast routes constitute loop-free data transmission paths from a data source to multiple receivers, namely, a multicast distribution tree.

In the ASM model, multicast routes include intra-domain routes and inter-domain routes.

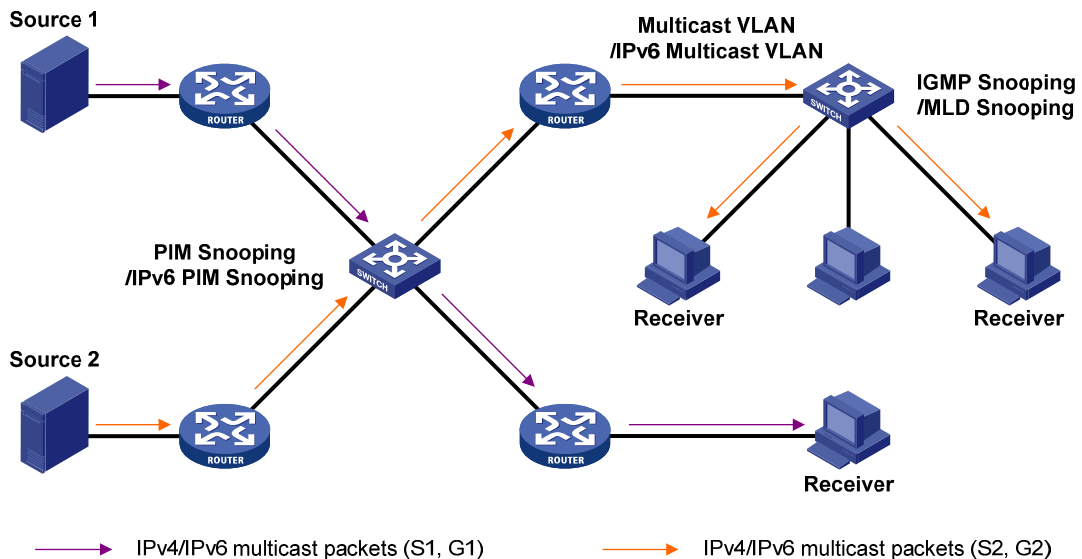
- An intra-domain multicast routing protocol discovers multicast sources and builds multicast distribution trees within an AS to deliver multicast data to receivers. Among a variety of mature intra-domain multicast routing protocols, Protocol Independent Multicast (PIM) is most widely used. Based on the forwarding mechanism, PIM has dense mode (often referred to as "PIM-DM"), and sparse mode (often referred to as "PIM-SM").
- An inter-domain multicast routing protocol is used for delivery of multicast information between two ASs. So far, mature solutions include Multicast Source Discovery Protocol (MSDP) and Multicast Border Gateway Protocol (MBGP). MSDP propagates multicast source information among different ASs. MBGP is an extension of the Multiprotocol Border Gateway Protocol (MP-BGP) for exchanging multicast routing information among different ASs.

For the SSM model, multicast routes are not divided into intra-domain routes and inter-domain routes. Because receivers know the position of the multicast source, channels established through PIM-SM are sufficient for the transport of multicast information.

Layer 2 multicast protocols

Layer 2 multicast protocols include IGMP snooping, MLD snooping, PIM snooping, IPv6 PIM snooping, multicast VLAN, and IPv6 multicast VLAN.

Figure 9 Positions of Layer 2 multicast protocols



- IGMP snooping and MLD snooping

IGMP snooping and MLD snooping are multicast constraining mechanisms that run on Layer 2 devices. They manage and control multicast groups by monitoring and analyzing IGMP or MLD messages exchanged between the hosts and Layer 3 multicast devices, effectively controlling the flooding of multicast data in a Layer 2 network.

- PIM snooping and IPv6 PIM snooping
PIM snooping and IPv6 PIM snooping run on Layer 2 devices. They determine which ports are interested in multicast data by analyzing the received IPv6 PIM messages, and add the ports to a multicast forwarding entry to make sure that multicast data can be forwarded to only the ports that are interested in the data.
- Multicast VLAN and IPv6 multicast VLAN
In the traditional multicast-on-demand mode, when users in different VLANs on a Layer 2 device need multicast information, the upstream Layer 3 device must forward a separate copy of the multicast data to each VLAN of the Layer 2 device. When the multicast VLAN or IPv6 multicast VLAN feature is enabled on the Layer 2 device, the Layer 3 multicast device sends only one copy of multicast to the multicast VLAN or IPv6 multicast VLAN on the Layer 2 device. This approach avoids waste of network bandwidth and extra burden on the Layer 3 device.

Multicast packet forwarding mechanism

In a multicast model, a multicast source sends information to the host group identified by the multicast group address in the destination address field of IP multicast packets. To deliver multicast packets to receivers located at different positions of the network, multicast routers on the forwarding paths usually need to forward multicast packets that an incoming interface receives to multiple outgoing interfaces. Compared with a unicast model, a multicast model is more complex in the following aspects:

- To ensure multicast packet transmission in the network, unicast routing tables or multicast routing tables (for example, the MBGP routing table) specially provided for multicast must be used as guidance for multicast forwarding.
- To process the same multicast information from different peers received on different interfaces of the same device, every multicast packet undergoes a reverse path forwarding (RPF) check on the incoming interface. The result of the RPF check determines whether the packet will be forwarded or discarded. The RPF check mechanism is the basis for most multicast routing protocols to implement multicast forwarding.

For more information about the RPF mechanism, see "[Configuring multicast routing and forwarding \(available only on the HP 5500 EI\)](#)" and "[Configuring IPv6 multicast routing and forwarding \(available only on the HP 5500 EI\)](#)."

Multicast support for VPNs

Multicast support for VPNs refers to multicast applied in virtual private networks (VPNs).

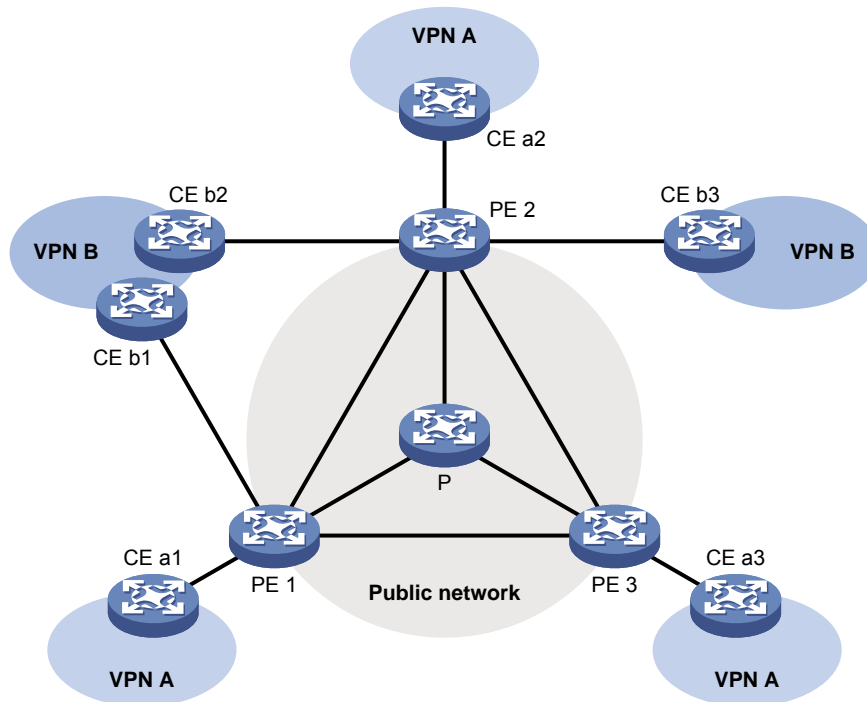
NOTE:

- Multicast support for VPNs is not available in IPv6 networks.
 - Multicast supporting for VPNs is not available for the HP 5500 SI switches.
-

Introduction to VPN instances

VPNs must be isolated from one another and from the public network. As shown in [Figure 10](#), VPN A and VPN B separately access the public network through PE devices.

Figure 10 VPN networking diagram



- The provider (P) device belongs to the public network. The customer edge (CE) devices belong to their respective VPNs. Each CE device serves its own VPN and maintains only one set of forwarding mechanisms.
- The provider edge (PE) devices connect to the public network and the VPNs at the same time. Each PE device must strictly distinguish the information for different networks, and maintain a separate forwarding mechanism for each network. On a PE device, a set of software and hardware that serve the same network forms an instance. Multiple instances can exist on the same PE device, and an instance can reside on different PE devices. On a PE device, the instance for the public network is called the public network instance, and those for VPNs are called VPN instances.

Multicast application in VPNs

A PE device that supports multicast for VPNs does the following operations:

- Maintains an independent set of independent multicast forwarding mechanisms for each VPN, including the multicast protocols, PIM neighbor information, and multicast routing table. In a VPN the device forwards multicast data based on the forwarding table or routing table for that VPN.
- Implements the isolation between different VPNs.
- Implements information exchange and data conversion between the public network and VPN instances.

As shown in [Figure 10](#), when a multicast source in VPN A sends a multicast stream to a multicast group, only the receivers that belong to both the multicast group and VPN A can receive the multicast stream. The multicast data is multicast both in VPN A and in the public network.

Configuring IGMP snooping

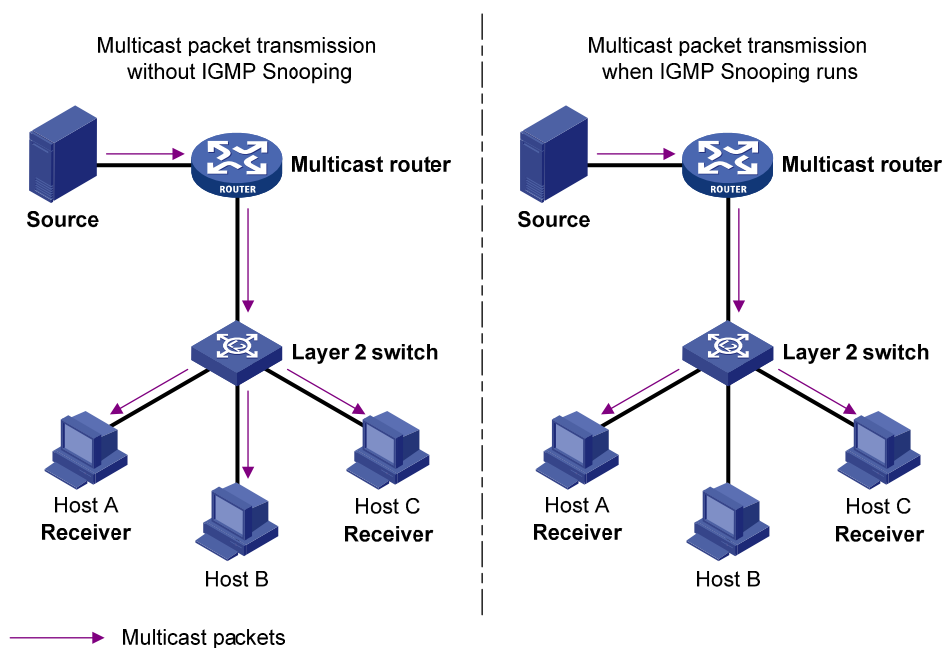
Overview

Internet Group Management Protocol (IGMP) snooping is a multicast constraining mechanism that runs on Layer 2 devices to manage and control multicast groups.

By analyzing received IGMP messages, a Layer 2 device that runs IGMP snooping establishes mappings between ports and multicast MAC addresses, and forwards multicast data based on these mappings.

As shown in [Figure 11](#), when IGMP snooping does not run on the Layer 2 switch, multicast packets are flooded to all devices at Layer 2. When IGMP snooping runs on the Layer 2 switch, multicast packets for known multicast groups are multicast to the receivers, rather than flooded to all hosts at Layer 2.

Figure 11 Before and after IGMP snooping is enabled on the Layer 2 device



IGMP snooping enables the Layer 2 switch to forward multicast data to only the receivers that require the data at Layer 2. It has the following advantages:

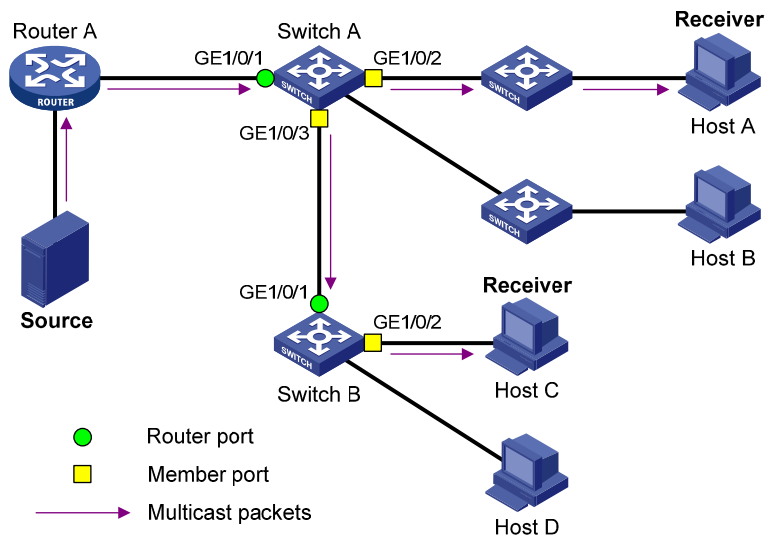
- Reducing Layer 2 broadcast packets, thus saving network bandwidth
- Enhancing the security of multicast traffic
- Facilitating the implementation of per-host accounting

Basic concepts in IGMP snooping

IGMP snooping related ports

As shown in [Figure 12](#), Router A connects to the multicast source, IGMP snooping runs on Switch A and Switch B, and Host A and Host C are receiver hosts (namely, members of a multicast group).

Figure 12 IGMP snooping related ports



Ports involved in IGMP snooping, as shown in [Figure 12](#), are described as follows:

- Router port**—A router port is a port on an Ethernet switch that leads the switch toward a Layer 3 multicast device (designated router or IGMP querier). In the figure, GigabitEthernet 1/0/1 of Switch A and GigabitEthernet 1/0/1 of Switch B are router ports. The switch registers all its local router ports in its router port list.

In this document, a router port is a port on a switch that leads the switch toward a Layer 3 multicast device. It is not a port on an ordinary router.

- Member port**—A member port is a port on an Ethernet switch that leads the switch toward multicast group members. In the figure, GigabitEthernet 1/0/2 and GigabitEthernet 1/0/3 of Switch A and GigabitEthernet 1/0/2 of Switch B are member ports. The switch registers all the member ports on the local device in its IGMP snooping forwarding table.

Unless otherwise specified, router ports and member ports in this document include both static and dynamic router ports and member ports.

NOTE:

An IGMP-snooping-enabled switch deems that all its ports on which IGMP general queries with the source IP address other than 0.0.0.0 or that receive PIM hello messages are received are dynamic router ports. For more information about PIM hello messages, see "[Configuring PIM \(available only on the HP 5500 EI\)](#)"

Aging timers for dynamic ports in IGMP snooping and related messages and actions

Timer	Description	Message before expiry	Action after expiry
Dynamic router port aging timer	For each dynamic router port, the switch sets an aging timer. When the timer expires, the dynamic router port ages out.	IGMP general query of which the source address is not 0.0.0.0 or PIM hello	The switch removes this port from its router port list.

Timer	Description	Message before expiry	Action after expiry
Dynamic member port aging timer	When a port dynamically joins a multicast group, the switch starts an aging timer for the port. When the timer expires, the dynamic member port ages out.	IGMP membership report	The switch removes this port from the IGMP snooping forwarding table.

NOTE:

In IGMP snooping, only dynamic ports age out. Static ports never age out.

How IGMP snooping works

In this section, the involved ports are dynamic ports. For information about how to configure and remove static ports, see "[Configuring static ports.](#)"

A switch that runs IGMP snooping performs different actions when it receives different IGMP messages.

When receiving a general query

The IGMP querier periodically sends IGMP general queries to all hosts and routers (224.0.0.1) on the local subnet to determine whether any active multicast group members exist on the subnet.

After receiving an IGMP general query, the switch forwards it to all ports in the VLAN (except the port that received the query). The switch also performs the following judgment:

- If the port that received the query is a dynamic router port in the router port list of the switch, the switch restarts the aging timer for the port.
- If the port is not in its router port list, the switch adds it into its router port list as a dynamic router port and starts an aging timer for the port.

When receiving a membership report

A host sends an IGMP report to the IGMP querier in the following circumstances:

- If the host has been a member of a multicast group, after receiving an IGMP query, the host responds to the query with an IGMP report.
- When the host wants to join a multicast group, it sends an IGMP report to the IGMP querier, specifying the multicast group to join.

After receiving an IGMP report, the switch forwards it through all the router ports in the VLAN, resolves the address of the reported multicast group. The switch also performs the following judgment:

- If no forwarding entry matches the group address, the switch creates a forwarding entry for the group, adds the port that received the IGMP report as a dynamic member port to the forwarding entry, and starts an aging timer for the port.
- If a forwarding entry matches the group address, but the port that received the IGMP report is not in the forwarding entry for the group, the switch adds the port as a dynamic member port to the forwarding entry, and starts an aging timer for the port.
- If a forwarding entry matches the group address and the port that received the IGMP report is in the forwarding entry for the group, the switch restarts the aging timer for the port.

A switch does not forward an IGMP report through a non-router port. The reason is that if the switch forwards a report message through a member port, all the attached hosts that are monitoring the

reported multicast address suppress their own reports after receiving this report according to the IGMP report suppression mechanism. This prevents the switch from confirming whether the reported multicast group still has active members attached to that port. For more information about the IGMP report suppression mechanism, see "[Configuring IGMP \(available only on the HP 5500 E1\)](#)"

When receiving a leave message

When an IGMPv1 host leaves a multicast group, the host does not send an IGMP leave message, and the switch cannot know immediately that the host has left the multicast group. However, because the host stops sending IGMP reports as soon as it leaves the multicast group, the switch removes the port that connects to the host from the forwarding entry for the multicast group when the aging timer for the port expires.

When an IGMPv2 or IGMPv3 host leaves a multicast group, the host sends an IGMP leave message to the multicast router.

When the switch receives an IGMP leave message on a dynamic member port, the switch first checks whether a forwarding entry matches the group address in the message, and, if a match is found, whether the forwarding entry for the group contains the dynamic member port.

- If no forwarding entry matches the group address, or if the forwarding entry does not contain the port, the switch directly discards the IGMP leave message.
- If a forwarding entry matches the group address and the forwarding entry contains the port, the switch forwards the leave message to all router ports in the VLAN. Because the switch does not know whether any other hosts attached to the port are still listening to that group address, the switch does not immediately remove the port from the forwarding entry for that group. Instead, it restarts the aging timer for the port.

After receiving the IGMP leave message, the IGMP querier resolves the multicast group address in the message and sends an IGMP group-specific query to the multicast group through the port that received the leave message. After receiving the IGMP group-specific query, the switch forwards it through all its router ports in the VLAN and all member ports of the multicast group. The switch also performs the following judgment for the port that received the IGMP leave message:

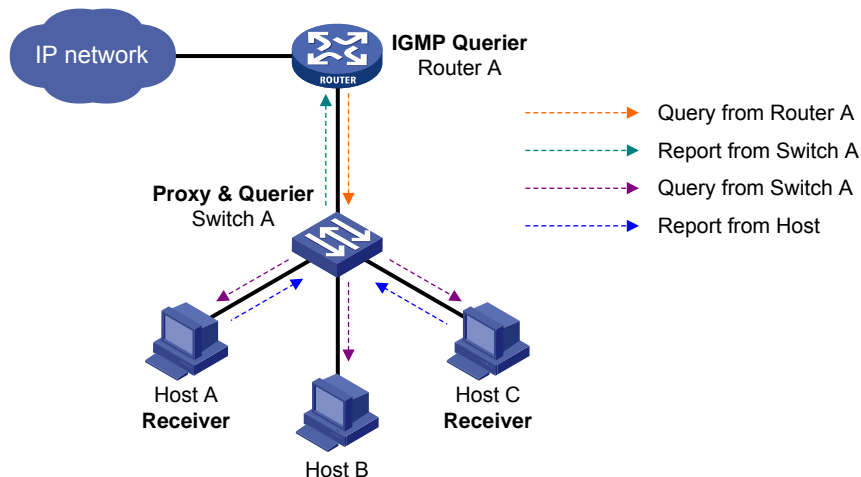
- If the port (assuming that it is a dynamic member port) receives an IGMP report in response to the group-specific query before its aging timer expires, it indicates that some host attached to the port is receiving or expecting to receive multicast data for the multicast group. The switch restarts the aging timer for the port.
- If the port receives no IGMP report in response to the group-specific query before its aging timer expires, it indicates that no hosts attached to the port are still listening to that group address. The switch removes the port from the forwarding entry for the multicast group when the aging timer expires.

IGMP snooping proxying

You can configure the IGMP snooping proxying function on an edge device to reduce the number of IGMP reports and leave messages sent to its upstream device. The device configured with IGMP snooping proxying is called an IGMP snooping proxy. It is a host from the perspective of its upstream device.

Even though an IGMP snooping proxy is a host from the perspective of its upstream device, the IGMP membership report suppression mechanism for hosts does not take effect on it. For more information about the IGMP report suppression mechanism for hosts, see "[Configuring IGMP \(available only on the HP 5500 E1\)](#)."

Figure 13 Network diagram



As shown in Figure 13, Switch A works as an IGMP snooping proxy. As a host from the perspective of the querier Router A, Switch A represents its attached hosts to send membership reports and leave messages to Router A.

Table 6 IGMP message processing on an IGMP snooping proxy

IGMP message	Actions
General query	When receiving an IGMP general query, the proxy forwards it to all ports but the receiving port. In addition, the proxy generates a report according to the group memberships it maintains and sends the report out of all router ports.
Group-specific query	In response to the IGMP group-specific query for a certain multicast group, the proxy sends the report to the group out of all router ports if the forwarding entry for the group still contains a member port.
Report	<p>After receiving a report for a multicast group, the proxy looks up the multicast forwarding table for the forwarding entry for the multicast group.</p> <ul style="list-style-type: none"> • If a forwarding entry matches the multicast group and contains the receiving port as a dynamic member port, the proxy restarts the aging timer for the port. • If a forwarding entry matches the multicast group but does not contain the receiving port, the proxy adds the port to the forwarding entry as a dynamic member port and starts an aging timer for the port. • If no forwarding entry matches the multicast group, the proxy creates a forwarding entry for the multicast group, adds the receiving port to the forwarding entry as a dynamic member port, and starts an aging timer for the port.
Leave	In response to an IGMP leave message for a multicast group, the proxy sends a group-specific query out of the receiving port. After making sure that no member port is contained in the forwarding entry for the multicast group, the proxy sends a leave message to the group out of all router ports.

Protocols and standards

RFC 4541, *Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches*

IGMP snooping configuration task list

Task	Remarks	
Configuring basic IGMP snooping functions	Enabling IGMP snooping	Required
	Specifying the version of IGMP snooping	Optional
	Configuring static multicast MAC address entries	Optional
Configuring IGMP snooping port functions	Setting aging timers for dynamic ports	Optional
	Configuring static ports	Optional
	Configuring a port as a simulated member host	Optional
	Enabling fast-leave processing	Optional
Configuring IGMP snooping querier	Disabling a port from becoming a dynamic router port	Optional
	Enabling IGMP snooping querier	Optional
	Configuring parameters for IGMP queries and responses	Optional
	Configuring the source IP addresses for IGMP queries	Optional
Configuring IGMP snooping proxying	Enabling IGMP snooping proxying	Optional
	Configuring a source IP address for the IGMP messages sent by the proxy	Optional
Configuring an IGMP snooping policy	Configuring a multicast group filter	Optional
	Configuring multicast source port filtering	Optional
	Enabling dropping unknown multicast data	Optional
	Configuring IGMP report suppression	Optional
	Setting the maximum number of multicast groups that a port can join	Optional
	Setting the 802.1p precedence for IGMP messages	Optional
	Enabling multicast group replacement	Optional
	Configuring a multicast user control policy	Optional
	Enabling the IGMP snooping host tracking function	Optional
Setting the DSCP value for IGMP messages	Optional	

For the configuration tasks in this section:

- In IGMP snooping view, configurations that you make are effective in all VLANs. In VLAN view, configurations that you make are effective on only the ports that belong to the current VLAN. For a given VLAN, a configuration that you make in IGMP snooping view is effective only if you do not make the same configuration in VLAN view.

- In IGMP snooping view, configurations that you make are effective on all ports. In Layer 2 Ethernet interface view or Layer 2 aggregate interface view, configurations that you make are effective only on the current port. In port group view, configurations that you make are effective on all ports in the current port group. For a given port, a configuration that you make in IGMP snooping view is effective only if you do not make the same configuration in Layer 2 Ethernet interface view, Layer 2 aggregate interface view, or port group view.
- For IGMP snooping, configurations that you make on a Layer 2 aggregate interface do not interfere with configurations that you make on its member ports, nor do they participate in aggregation calculations. Configurations that you make on a member port of an aggregate group do not take effect until it leaves the aggregate group.

Configuring basic IGMP snooping functions

Before you configure basic IGMP snooping functions, complete the following tasks:

- Configure the corresponding VLANs.
- Determine the version of IGMP snooping.

Enabling IGMP snooping

Configuration guidelines

- You must enable IGMP snooping globally before you enable it in a VLAN.
- After you enable IGMP snooping in a VLAN, do not enable IGMP or PIM on the corresponding VLAN interface.
- When you enable IGMP snooping in a specified VLAN, IGMP snooping works only on the ports in this VLAN.

Configuration procedure

To enable IGMP snooping:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IGMP snooping globally and enter IGMP-snooping view.	igmp-snooping	Disabled by default
3. Return to system view.	quit	N/A
4. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
5. Enable IGMP snooping in the VLAN.	igmp-snooping enable	Disabled by default

Specifying the version of IGMP snooping

Configuration guidelines

Different versions of IGMP snooping can process different versions of IGMP messages:

- IGMPv2 snooping can process IGMPv1 and IGMPv2 messages, but cannot process IGMPv3 messages, which will be flooded in the VLAN.
- IGMPv3 snooping can process IGMPv1, IGMPv2 and IGMPv3 messages.

If you change IGMPv3 snooping to IGMPv2 snooping, the system clears all IGMP snooping forwarding entries that are dynamically added, and also does the following:

- Keeps static IGMPv3 snooping forwarding entries (*, G).
- Clears static IGMPv3 snooping forwarding entries (S, G), which will be restored when IGMP snooping is switched back to IGMPv3 snooping.

For more information about static joins, see "[Configuring static ports.](#)"

Configuration procedure

To specify the version of IGMP snooping:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
3. Specify the version of IGMP snooping.	igmp-snooping version <i>version-number</i>	Version 2 by default

Configuring static multicast MAC address entries

Configuration guidelines

In Layer-2 multicast, a Layer 2 multicast protocol (such as, IGMP snooping) can dynamically add multicast MAC address entries. Or, you can manually configure multicast MAC address entries.

In system view, the configuration is effective for the specified ports. In interface view or port group view, the configuration is effective only on the current port or the ports in the current port group.

Any legal multicast MAC address except 0100-5Exx-xxxx (where x represents a hexadecimal number from 0 to F) can be manually added to the multicast MAC address table. Multicast MAC addresses are the MAC addresses whose the least significant bit of the most significant octet is 1.

Configuration procedure

To configure a static multicast MAC address entry in system view:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure a static multicast MAC address entry.	mac-address multicast <i>mac-address interface interface-list</i> vlan <i>vlan-id</i>	No static multicast MAC address entries exist by default.

To configure static multicast MAC address entries in interface view:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter Layer 2 Ethernet interface view, Layer 2 aggregate interface view, or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Configure a static multicast MAC address entry.	mac-address multicast <i>mac-address</i> vlan <i>vlan-id</i>	No static multicast MAC address entries exist by default.

Configuring IGMP snooping port functions

Configuration prerequisites

Before you configure IGMP snooping port functions, complete the following tasks:

- Enable IGMP snooping in the VLAN.
- Configure the corresponding port groups.
- Determine the aging time of dynamic router ports.
- Determine the aging time of dynamic member ports.
- Determine the multicast group and multicast source addresses.

Setting aging timers for dynamic ports

If a switch receives no IGMP general queries or PIM hello messages on a dynamic router port when the aging timer of the port expires, the switch removes the port from the router port list.

If the switch receives no IGMP reports for a multicast group on a dynamic member port when the aging timer of the port expires, the switch removes the port from the multicast forwarding entry for that multicast group.

If the memberships of multicast groups change frequently, you can set a relatively small value for the aging timer of the dynamic member ports. If the memberships of multicast groups change rarely, you can set a relatively large value.

Configuring aging timers for dynamic ports globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IGMP-snooping view.	igmp-snooping	N/A
3. Set the aging timer for dynamic router ports.	router-aging-time <i>interval</i>	105 seconds by default
4. Set the aging timer for dynamic member ports.	host-aging-time <i>interval</i>	260 seconds by default

Configuring aging timers for dynamic ports in a VLAN

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
3. Set the aging timer for dynamic router ports.	igmp-snooping router-aging-time <i>interval</i>	105 seconds by default
4. Set the aging timer for dynamic member ports.	igmp-snooping host-aging-time <i>interval</i>	260 seconds by default

Configuring static ports

Configuration guidelines

If all hosts attached to a port are interested in the multicast data addressed to a particular multicast group or the multicast data that a particular multicast source sends to a particular group, you can configure the port as a static member port for the specified multicast group or the specified multicast source and group.

You can also configure a port as a static router port, through which the switch can forward all the multicast traffic that it received.

A static member port does not respond to queries from the IGMP querier; when you configure a port as a static member port or cancel this configuration on the port, the port does not send an unsolicited IGMP report or an IGMP leave message.

Static member ports and static router ports never age out. To remove such a port, use the corresponding **undo** command.

Configuration procedure

To configure static ports:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet interface view, Layer 2 aggregate interface view, or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Configure the port as a static member port.	igmp-snooping static-group <i>group-address</i> [source-ip <i>source-address</i>] vlan <i>vlan-id</i>	No static member ports exist by default.
4. Configure the port as a static router port.	igmp-snooping static-router-port vlan <i>vlan-id</i>	No static router ports exist by default.

Configuring a port as a simulated member host

Configuration guidelines

Generally, a host that runs IGMP can respond to IGMP queries that the IGMP querier sends. If a host fails to respond, the multicast router might deem that no member of this multicast group exists on the network segment, and removes the corresponding forwarding path.

To avoid this situation, you can configure the port as a simulated member host for a multicast group. A simulated host is equivalent to an independent host. For example, when a simulated member host receives an IGMP query, it gives a response separately. Therefore, the switch can continue receiving multicast data.

A simulated host acts like a real host in the following ways:

- When a port is configured as a simulated member host, the switch sends an unsolicited IGMP report through the port, and can respond to IGMP general queries with IGMP reports through the port.
- When the simulated joining function is disabled on a port, the switch sends an IGMP leave message through the port.

Unlike a static member port, a port that you configure as a simulated member host ages out like a dynamic member port.

Configuration procedure

To configure a port as a simulated member host:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet interface view, Layer 2 aggregate interface view, or port group view.	<ul style="list-style-type: none">• Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i>• Enter port group view: port-group manual <i>port-group-name</i>	Use either command.
3. Configure a port as a simulated member host.	igmp-snooping host-join <i>group-address</i> [source-ip <i>source-address</i>] vlan <i>vlan-id</i>	Not configured by default.

Enabling fast-leave processing

Configuration guidelines

The fast-leave processing feature enables the switch to process IGMP leave messages quickly. With the fast-leave processing feature enabled, when the switch receives an IGMP leave message on a port, it immediately removes that port from the forwarding entry for the multicast group specified in the message. Then, when the switch receives IGMP group-specific queries for that multicast group, it does not forward them to that port.

On a port that has only one host attached, you can enable fast-leave processing to save bandwidth and resources. However, on a port that has multiple hosts attached, you should not enable fast-leave

processing if you have enabled dropping unknown multicast data globally or for the port. Otherwise, if a host on the port leaves a multicast group, the other hosts attached to the port in the same multicast group cannot receive the multicast data for the group.

Configuration procedure

To enable fast-leave processing globally:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IGMP-snooping view.	igmp-snooping	N/A
3. Enable fast-leave processing.	fast-leave [vlan vlan-list]	Disabled by default

To enable fast-leave processing for a port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet interface view, Layer 2 aggregate interface view, or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface interface-type interface-number Enter port group view: port-group manual port-group-name 	Use either command.
3. Enable fast-leave processing.	igmp-snooping fast-leave [vlan vlan-list]	Disabled by default.

Disabling a port from becoming a dynamic router port

The following problems might exist in a multicast access network:

- After receiving an IGMP general query or a PIM hello message from a connected host, a router port becomes a dynamic router port. Before its timer expires, this dynamic router port receives all multicast packets within the VLAN where the port belongs, and forwards them to the host, affecting normal multicast reception of the host.
- In addition, the IGMP general query or PIM hello message that the host sends affects the multicast routing protocol state on Layer 3 devices, such as the IGMP querier or DR election, and might further cause network interruption.

To solve these problems, disable that router port from becoming a dynamic router port after the port receives an IGMP general query or a PIM hello message, so as to improve network security and control over multicast users.

To disable a port from becoming a dynamic router port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter Layer 2 Ethernet interface view, Layer 2 aggregate interface view, or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Disable the ports from becoming dynamic router port.	igmp-snooping router-port-deny [vlan <i>vlan-list</i>]	By default, a port can become a dynamic router port.

NOTE:

This configuration does not affect the static router port configuration.

Configuring IGMP snooping querier

Configuration prerequisites

Before you configure IGMP snooping querier, complete the following tasks:

- Enable IGMP snooping in the VLAN.
- Determine the IGMP general query interval.
- Determine the IGMP last-member query interval.
- Determine the maximum response delay for IGMP general queries.
- Determine the source address of IGMP general queries.
- Determine the source address of IGMP group-specific queries.

Enabling IGMP snooping querier

In an IP multicast network that runs IGMP, a multicast router or Layer 3 multicast switch sends IGMP queries, so that all Layer 3 multicast devices can establish and maintain multicast forwarding entries, in order to forward multicast traffic correctly at the network layer. This router or Layer 3 switch is called the "IGMP querier". For more information about IGMP querier, see "[Configuring IGMP \(available only on the HP 5500 EI\)](#)."

However, a Layer 2 multicast switch does not support IGMP, and therefore cannot send general queries by default. When you enable IGMP snooping querier on a Layer 2 switch in a VLAN where multicast traffic is switched only at Layer 2 and no multicast routers are present, the Layer 2 switch sends IGMP queries, so that multicast forwarding entries can be established and maintained at the data link layer.

To enable IGMP snooping querier:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A

Step	Command	Remarks
3. Enable IGMP snooping querier.	igmp-snooping querier	Disabled by default

! **IMPORTANT:**

In a multicast network that runs IGMP, you do not need to configure an IGMP snooping querier because it may affect IGMP querier elections by sending IGMP general queries with a low source IP address.

Configuring parameters for IGMP queries and responses

Configuration guidelines

You can modify the IGMP general query interval based on actual condition of the network.

A multicast listening host starts a timer for each multicast group that it has joined when it receives an IGMP query (general query or group-specific query). This timer is initialized to a random value in the range of 0 to the maximum response delay advertised in the IGMP query message. When the timer value decreases to 0, the host sends an IGMP report to the multicast group.

To speed up the response of hosts to IGMP queries and avoid simultaneous timer expirations causing IGMP report traffic bursts, you must properly set the maximum response delay.

- The maximum response delay for IGMP general queries is set by the **max-response-time** command.
- The maximum response delay for IGMP group-specific queries equals the IGMP last-member query interval.

In the configuration, make sure that the IGMP general query interval is larger than the maximum response delay for IGMP general queries. Otherwise, multicast group members might be deleted by mistake.

Configuration procedure

To configure the global parameters for IGMP queries and responses:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IGMP-snooping view.	igmp-snooping	N/A
3. Set the maximum response delay for IGMP general queries.	max-response-time <i>interval</i>	10 seconds by default
4. Set the IGMP last-member query interval.	last-member-query-interval <i>interval</i>	1 second by default

To configure the parameters for IGMP queries and responses in a VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
3. Set the interval for sending IGMP general queries.	igmp-snooping query-interval <i>interval</i>	60 seconds by default

Step	Command	Remarks
4.	Set the maximum response delay for IGMP general queries.	igmp-snooping max-response-time <i>interval</i> 10 seconds by default
5.	Set the IGMP last-member query interval.	igmp-snooping last-member-query-interval <i>interval</i> 1 second by default

Configuring the source IP addresses for IGMP queries

After the switch receives an IGMP query whose source IP address is 0.0.0.0 on a port, it does not enlist that port as a dynamic router port. This might prevent multicast forwarding entries from being correctly created at the data link layer and eventually cause multicast traffic forwarding to fail. To avoid this problem, when a Layer 2 switch acts as the IGMP snooping querier, HP recommends you to configure a non-all-zero IP address as the source IP address of IGMP queries.

ⓘ IMPORTANT:

The source address of IGMP query messages might affect the IGMP querier election within the segment

To configure the source IP addresses for IGMP queries:

Step	Command	Remarks
1.	Enter system view.	system-view N/A
2.	Enter VLAN view.	vlan <i>vlan-id</i> N/A
3.	Configure the source address of IGMP general queries.	igmp-snooping general-query source-ip { <i>ip-address</i> current-interface } 0.0.0.0 by default
4.	Configure the source IP address of IGMP group-specific queries.	igmp-snooping special-query source-ip { <i>ip-address</i> current-interface } 0.0.0.0 by default

Configuring IGMP snooping proxying

Configuration prerequisites

Before you configure IGMP snooping proxying in a VLAN, complete the following tasks:

- Enable IGMP snooping in the VLAN.
- Determine the source IP address for the IGMP reports sent by the proxy.
- Determine the source IP address for the IGMP leave messages sent by the proxy.

Enabling IGMP snooping proxying

The IGMP snooping proxying function works on a per-VLAN basis. After you enable the function in a VLAN, the device works as the IGMP snooping proxy for the downstream hosts and upstream router in the VLAN.

To enable IGMP snooping proxying in a VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
3. Enable IGMP snooping proxying in the VLAN.	igmp-snooping proxying enable	Disabled by default

Configuring a source IP address for the IGMP messages sent by the proxy

You can set the source IP addresses in the IGMP reports and leave messages that the IGMP snooping proxy sends on behalf of its attached hosts.

To configure the source IP addresses for the IGMP messages that the IGMP snooping proxy sends on behalf of its attached hosts in a VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
3. Configure a source IP address for the IGMP reports that the proxy sends.	igmp-snooping report source-ip { <i>ip-address</i> current-interface }	The default is 0.0.0.0.
4. Configure a source IP address for the IGMP leave messages that the proxy sends.	igmp-snooping leave source-ip { <i>ip-address</i> current-interface }	The default is 0.0.0.0.

Configuring an IGMP snooping policy

Configuration prerequisites

Before you configure an IGMP snooping policy, complete the following tasks:

- Enable IGMP snooping in the VLAN.
- Determine the ACL rule for multicast group filtering.
- Determine the maximum number of multicast groups that a port can join.
- Determine the 802.1p precedence for IGMP messages.

Configuring a multicast group filter

On an IGMP snooping-enabled switch, you can configure a multicast group filter to limit multicast programs available to users.

Configuration guidelines

In an application, when a user requests a multicast program, the user's host initiates an IGMP report. After receiving this report message, the switch resolves the multicast group address in the report and looks up the ACL. If a match is found to permit the port that received the report to join the multicast group,

the switch creates an IGMP snooping forwarding entry for the multicast group and adds the port to the forwarding entry. Otherwise, the switch drops this report message, in which case, the multicast data for the multicast group is not sent to this port, and the user cannot retrieve the program.

When you configure a multicast group filter in a multicast VLAN, be sure to configure the filter in the sub-VLANs of the multicast VLAN. Otherwise, the configuration does not take effect.

In IGMPv3, when a host is enabled to join multiple multicast groups, the multicast group filter cannot correctly filter multicast groups because the host that runs IGMPv3 sends multiple multicast groups that it wants to join in one membership report.

Configuration procedure

To configure a multicast group filter globally:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IGMP-snooping view.	igmp-snooping	N/A
3. Configure a multicast group filter.	group-policy <i>acl-number</i> [vlan <i>vlan-list</i>]	By default, no group filter is globally configured. That is, the hosts in a VLAN can join any valid multicast group.

To configure a multicast group filter for a port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet interface view, Layer 2 aggregate interface view, or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Configure a multicast group filter.	igmp-snooping group-policy <i>acl-number</i> [vlan <i>vlan-list</i>]	By default, no group filter is configured on the current port. That is, the hosts on this port can join any valid multicast group.

Configuring multicast source port filtering

When the multicast source port filtering feature is enabled on a port, the port can connect to only multicast receivers rather than to multicast sources, because the port blocks all multicast data packets but it permits multicast protocol packets to pass.

If this feature is disabled on a port, the port can connect to both multicast sources and multicast receivers.

Configuring multicast source port filtering globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IGMP-snooping view.	igmp-snooping	N/A
3. Enable multicast source port filtering.	source-deny port <i>interface-list</i>	Disabled by default

Configuring multicast source port filtering on a port

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet interface view or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Enable multicast source port filtering.	igmp-snooping source-deny	Disabled by default.

Enabling dropping unknown multicast data

Configuration guidelines

Unknown multicast data refers to multicast data for which no entries exist in the IGMP snooping forwarding table. When the switch receives such multicast traffic, one of the following occurs:

- When the function of dropping unknown multicast data is disabled, the switch floods unknown multicast data in the VLAN that the unknown multicast data belongs to, causing network bandwidth waste and low forwarding efficiency.
- When the function of dropping unknown multicast data is enabled, the switch forwards unknown multicast data to its router ports instead of flooding it in the VLAN. If no router ports exist, the switch drops the unknown multicast data.

Configuration procedure

To enable dropping unknown multicast data in a VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
3. Enable dropping unknown multicast data.	igmp-snooping drop-unknown	Disabled by default

Configuring IGMP report suppression

When a Layer 2 switch receives an IGMP report from a multicast group member, the switch forwards the message to the Layer 3 device that directly connects to the Layer 2 switch. When multiple members of a multicast group are attached to the Layer 2 switch, the Layer 3 device might receive duplicate IGMP reports for the multicast group from these members.

With the IGMP report suppression function enabled, within each query interval, the Layer 2 switch forwards only the first IGMP report for the multicast group to the Layer 3 device. It does not forward the subsequent IGMP reports for the same multicast group. This helps reduce the number of packets being transmitted over the network.

! IMPORTANT:

On an IGMP snooping proxy, IGMP membership reports are suppressed if the entries for the corresponding groups exist in the forwarding table, no matter the suppression function is enabled or not.

To configure IGMP report suppression:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IGMP-snooping view.	igmp-snooping	N/A
3. Enable IGMP report suppression.	report-aggregation	Enabled by default

Setting the maximum number of multicast groups that a port can join

To regulate multicast traffic on a port, configure the maximum number of multicast groups that the port can join.

When you configure this maximum number, if the number of multicast groups the port has joined exceeds the configured maximum value, the system deletes all the forwarding entries for the port from the IGMP snooping forwarding table, and the hosts on this port join multicast groups again until the number of multicast groups that the port joins reaches the maximum value. When the port joins a multicast group, if the port has been configured as a static member port, the system applies the configurations to the port again. If you have configured simulated joining on the port, the system establishes corresponding forwarding entry for the port after receiving a report from the simulated member host.

To set the maximum number of multicast groups that a port can join:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter Layer 2 Ethernet interface view, Layer 2 aggregate interface view, or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Set the maximum number of multicast groups that a port can join.	igmp-snooping group-limit <i>limit</i> [vlan <i>vlan-list</i>]	By default, the upper limit is 2000 for the HP 5500 EI switches, and 1000 for the HP 5500 SI switches.

Enabling multicast group replacement

Configuration guidelines

For various reasons, the number of multicast groups that the switch or a port joins might exceed the upper limit. In addition, in some specific applications, a multicast group that the switch newly joins must replace an existing multicast group automatically. A typical example is channel switching. To view a new channel, a user switches from the current multicast group to the new one.

To realize such requirements, you can enable the multicast group replacement function on the switch or on a certain port. When the number of multicast groups that the switch or on the port has joined reaches the limit, one of the following occurs:

- If the multicast group replacement feature is disabled, new IGMP reports are automatically discarded.
- If the multicast group replacement feature is enabled, the multicast group that the switch or a port newly joins automatically replaces an existing multicast group that has the lowest address.

In the configuration, be sure to configure the maximum number of multicast groups allowed on a port (see "[Setting the maximum number of multicast groups that a port can join](#)") before enabling multicast group replacement. Otherwise, the multicast group replacement functionality will not take effect.

Configuration procedure

To enable multicast group replacement globally:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IGMP-snooping view.	igmp-snooping	N/A
3. Enable multicast group replacement.	overflow-replace [vlan <i>vlan-list</i>]	Disabled by default

To enable multicast group replacement for a port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter Layer 2 Ethernet interface view, Layer 2 aggregate interface view, or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Enable multicast group replacement.	igmp-snooping overflow-replace [vlan <i>vlan-list</i>]	Disabled by default.

Setting the 802.1p precedence for IGMP messages

You can change the 802.1p precedence for IGMP messages so that they can be assigned higher forwarding priority when congestion occurs on their outgoing ports.

Setting the 802.1p precedence for IGMP messages globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IGMP-snooping view.	igmp-snooping	N/A
3. Set the 802.1p precedence for IGMP messages.	dot1p-priority <i>priority-number</i>	The default 802.1p precedence for IGMP messages is 0.

Setting the 802.1p precedence for IGMP messages in a VLAN

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
3. Set the 802.1p precedence for IGMP messages in the VLAN.	igmp-snooping dot1p-priority <i>priority-number</i>	The default 802.1p precedence for IGMP messages is 0.

Configuring a multicast user control policy

Configuration guidelines

Multicast user control policies are configured on access switches to allow only authorized users to receive requested multicast traffic flows. This helps restrict users from ordering certain multicast-on-demand programs.

In practice, a device first needs to perform authentication (802.1X authentication, for example) on connected hosts through a RADIUS server. Then, the device uses the configured multicast user control policy to perform multicast access control on authenticated users as follows:

- After receiving an IGMP report from a host, the access switch matches the multicast group address and multicast source address carried in the report with the configured policies. If a match is found,

the host is allowed to join the multicast group. Otherwise, the join report is dropped by the access switch.

- After receiving an IGMP leave message from a host, the access switch matches the multicast group and source addresses with the policies. If a match is found, the host is allowed to leave the group. Otherwise, the leave message is dropped by the access switch.

A multicast user control policy is functionally similar to a multicast group filter. A difference is that a control policy can control both multicast joining and leaving of users based on authentication and authorization, but a multicast group filter is configured on a port to control only multicast joining but not leaving of users without authentication or authorization.

Configuration procedure

To configure a multicast user control policy:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a user profile and enter its view.	user-profile <i>profile-name</i>	N/A
3. Configure a multicast user control policy.	igmp-snooping access-policy <i>acl-number</i>	No policy is configured by default. That is, a host can join or leave a valid multicast group at any time.
4. Return to system view.	quit	N/A
5. Enable the created user profile.	user-profile <i>profile-name</i> enable	Disabled by default.

For more information about the **user-profile** and **user-profile enable** commands, see *Security Command Reference*.

Enabling the IGMP snooping host tracking function

With the IGMP snooping host tracking function, the switch can record the information of the member hosts that are receiving multicast traffic, including the host IP address, running duration, and timeout time. You can monitor and manage the member hosts according to the recorded information.

Enabling the IGMP snooping host tracking function globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IGMP-snooping view.	igmp-snooping	N/A
3. Enable the IGMP snooping host tracking function globally.	host-tracking	Disabled by default

Enabling the IGMP snooping host tracking function in a VLAN

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
3. Enable the IGMP snooping host tracking function in the VLAN.	igmp-snooping host-tracking	Disabled by default

Setting the DSCP value for IGMP messages

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IGMP-snooping view.	igmp-snooping	N/A
3. Set the DSCP value for IGMP messages	dscp <i>dscp-value</i>	By default, the DSCP value in IGMP messages is 48.

NOTE:

This configuration applies to only the IGMP messages that the local switch generates rather than those forwarded ones.

Displaying and maintaining IGMP snooping

Task	Command	Remarks
Display IGMP snooping group information.	display igmp-snooping group [vlan <i>vlan-id</i>] [slot <i>slot-number</i>] [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about the hosts tracked by IGMP snooping.	display igmp-snooping host vlan <i>vlan-id group group-address</i> [source <i>source-address</i>] [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display static multicast MAC address entries.	display mac-address [<i>mac-address</i> [vlan <i>vlan-id</i>] [multicast] [vlan <i>vlan-id</i>] [count]] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display statistics for the IGMP messages learned by IGMP snooping.	display igmp-snooping statistics [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Remove all the dynamic group entries of a specified IGMP snooping group or all IGMP snooping groups.	reset igmp-snooping group { <i>group-address</i> all } [vlan <i>vlan-id</i>]	Available in user view
Clear statistics for the IGMP messages learned by IGMP snooping.	reset igmp-snooping statistics	Available in user view

NOTE:

- The **reset igmp-snooping group** command works only on an IGMP snooping-enabled VLAN, but not in a VLAN with IGMP enabled on its VLAN interface.
 - The **reset igmp-snooping group** command cannot remove the static group entries of IGMP snooping groups.
-

IGMP snooping configuration examples

Group policy and simulated joining configuration example

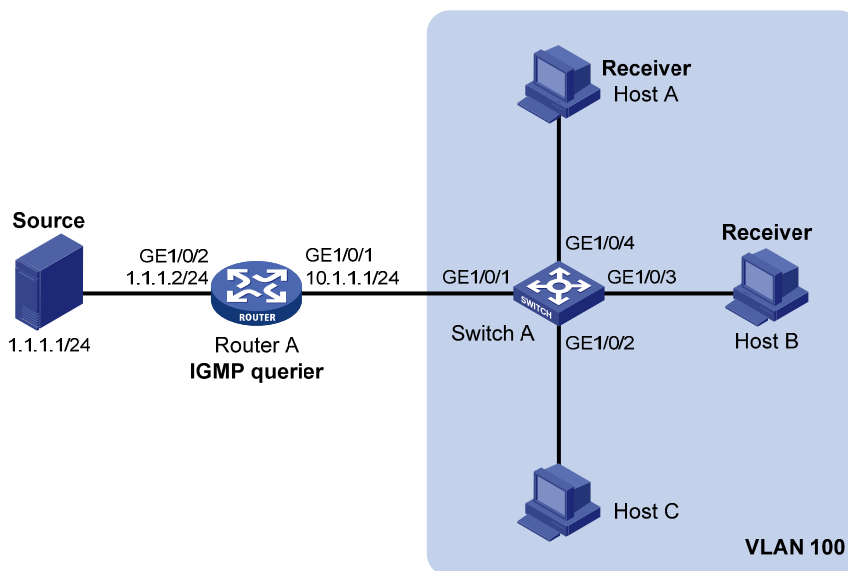
Network requirements

As shown in [Figure 14](#), IGMPv2 runs on Router A, IGMPv2 snooping runs on Switch A, and Router A acts as the IGMP querier on the subnet.

The receivers, Host A and Host B, can receive multicast traffic addressed to multicast group 224.1.1.1 only.

Multicast data for group 224.1.1.1 can be forwarded through GigabitEthernet 1/0/3 and GigabitEthernet 1/0/4 of Switch A even if Host A and Host B accidentally, temporarily stop receiving multicast data, and that Switch A drops unknown multicast data and does not broadcast the data to the VLAN where Switch A resides.

Figure 14 Network diagram



Configuration procedure

1. Configure IP addresses:
Configure an IP address and subnet mask for each interface as per [Figure 14](#). (Details not shown.)
2. Configure Router A:
Enable IP multicast routing, enable PIM-DM on each interface, and enable IGMP on GigabitEthernet 1/0/1.

```
<RouterA> system-view  
[RouterA] multicast routing-enable
```

```
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] igmp enable
[RouterA-GigabitEthernet1/0/1] pim dm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] pim dm
[RouterA-GigabitEthernet1/0/2] quit
```

3. Configure Switch A:

Enable IGMP snooping globally.

```
<SwitchA> system-view
[SwitchA] igmp-snooping
[SwitchA-igmp-snooping] quit
```

Create VLAN 100, assign GigabitEthernet 1/0/1 through GigabitEthernet 1/0/4 to this VLAN, and enable IGMP snooping and the function of dropping unknown multicast traffic in the VLAN.

```
[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/4
[SwitchA-vlan100] igmp-snooping enable
[SwitchA-vlan100] igmp-snooping drop-unknown
[SwitchA-vlan100] quit
```

Configure a multicast group filter so that the hosts in VLAN 100 can join only the multicast group 224.1.1.1.

```
[SwitchA] acl number 2001
[SwitchA-acl-basic-2001] rule permit source 224.1.1.1 0
[SwitchA-acl-basic-2001] quit
[SwitchA] igmp-snooping
[SwitchA-igmp-snooping] group-policy 2001 vlan 100
[SwitchA-igmp-snooping] quit
```

Configure GigabitEthernet 1/0/3 and GigabitEthernet 1/0/4 as simulated hosts for multicast group 224.1.1.1.

```
[SwitchA] interface gigabitethernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] igmp-snooping host-join 224.1.1.1 vlan 100
[SwitchA-GigabitEthernet1/0/3] quit
[SwitchA] interface gigabitethernet 1/0/4
[SwitchA-GigabitEthernet1/0/4] igmp-snooping host-join 224.1.1.1 vlan 100
[SwitchA-GigabitEthernet1/0/4] quit
```

4. Verify the configuration:

Display detailed IGMP snooping groups information in VLAN 100 on Switch A.

```
[SwitchA] display igmp-snooping group vlan 100 verbose
Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).
```

Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port

Subvlan flags: R-Real VLAN, C-Copy VLAN

Vlan(id):100.

Total 1 IP Group(s).

Total 1 IP Source(s).

```

Total 1 MAC Group(s).
Router port(s):total 1 port.
    GE1/0/1                (D) ( 00:01:30 )
IP group(s):the following ip group(s) match to one mac group.
  IP group address:224.1.1.1
  (0.0.0.0, 224.1.1.1):
    Attribute:    Host Port
  Host port(s):total 2 port.
    GE1/0/3      (D) ( 00:03:23 )
    GE1/0/4      (D) ( 00:04:10 )
MAC group(s):
  MAC group address:0100-5e01-0101
  Host port(s):total 2 port.
    GE1/0/3
    GE1/0/4

```

The output shows that GigabitEthernet 1/0/3 and GigabitEthernet 1/0/4 of Switch A has joined multicast group 224.1.1.1.

Static port configuration example

Network requirements

As shown in [Figure 15](#), IGMPv2 runs on Router A, and IGMPv2 snooping runs on Switch A, Switch B, and Switch C. Router A acts as the IGMP querier.

Host A and host C are permanent receivers of multicast group 224.1.1.1. GigabitEthernet 1/0/3 and GigabitEthernet 1/0/5 on Switch C are required to be configured as static member ports for multicast group 224.1.1.1 to enhance the reliability of multicast traffic transmission.

Suppose STP runs on the network. To avoid data loops, the forwarding path from Switch A to Switch C is blocked under normal conditions, and multicast traffic flows to the receivers attached to Switch C only along the path of Switch A—Switch B—Switch C.

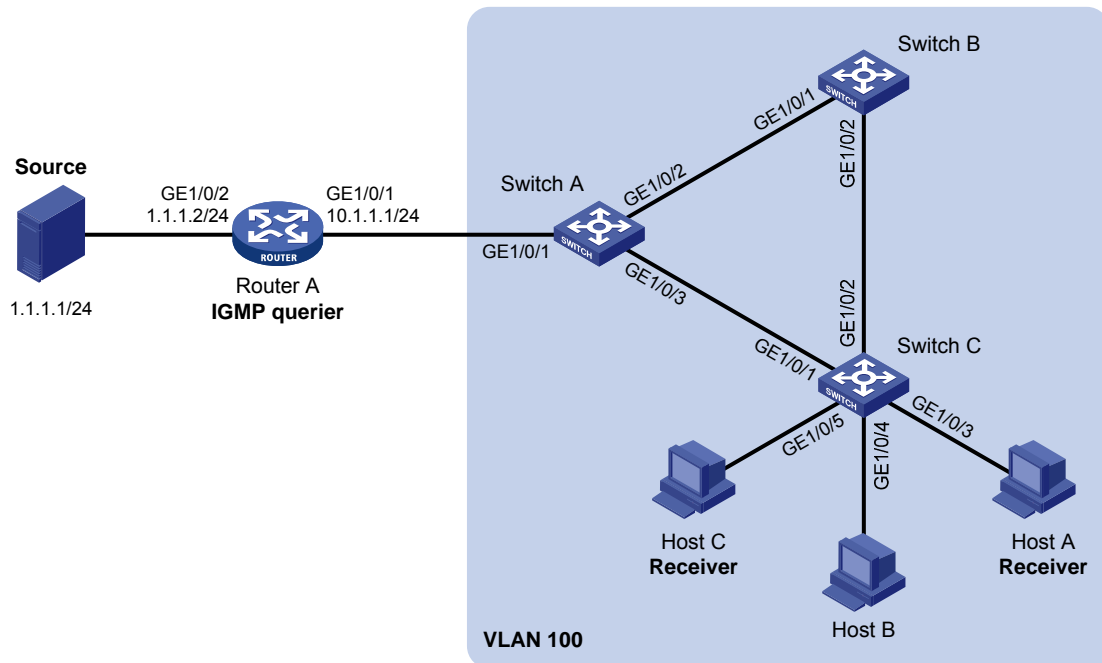
Configure GigabitEthernet 1/0/3 on Switch A as a static router port, so that multicast traffic can flow to the receivers nearly uninterrupted along the path of Switch A—Switch C in the case that the path of Switch A—Switch B—Switch C gets blocked.

For more information about the Spanning Tree Protocol (STP), see *Layer 2—LAN Switching Configuration Guide*.

NOTE:

If no static router port is configured, when the path of Switch A—Switch B—Switch C gets blocked, at least one IGMP query-response cycle must be completed before the multicast data can flow to the receivers along the new path of Switch A—Switch C. Namely multicast delivery will be interrupted during this process.

Figure 15 Network diagram



Configuration procedure

1. Configure IP addresses:

Configure an IP address and subnet mask for each interface as per Figure 15. (Details not shown.)

2. Configure Router A:

Enable IP multicast routing, enable PIM-DM on each interface, and enable IGMP on GigabitEthernet 1/0/1.

```
<RouterA> system-view
[RouterA] multicast routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] igmp enable
[RouterA-GigabitEthernet1/0/1] pim dm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] pim dm
[RouterA-GigabitEthernet1/0/2] quit
```

3. Configure Switch A:

Enable IGMP snooping globally.

```
<SwitchA> system-view
[SwitchA] igmp-snooping
[SwitchA-igmp-snooping] quit
```

Create VLAN 100, assign GigabitEthernet 1/0/1 through GigabitEthernet 1/0/3 to this VLAN, and enable IGMP snooping in the VLAN.

```
[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/3
[SwitchA-vlan100] igmp-snooping enable
[SwitchA-vlan100] quit
```

```
# Configure GigabitEthernet 1/0/3 to be a static router port.
```

```
[SwitchA] interface gigabitethernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] igmp-snooping static-router-port vlan 100
[SwitchA-GigabitEthernet1/0/3] quit
```

4. Configure Switch B:

```
# Enable IGMP snooping globally.
```

```
<SwitchB> system-view
[SwitchB] igmp-snooping
[SwitchB-igmp-snooping] quit
```

```
# Create VLAN 100, assign GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 to this VLAN,
and enable IGMP snooping in the VLAN.
```

```
[SwitchB] vlan 100
[SwitchB-vlan100] port gigabitethernet 1/0/1 gigabitethernet 1/0/2
[SwitchB-vlan100] igmp-snooping enable
[SwitchB-vlan100] quit
```

5. Configure Switch C:

```
# Enable IGMP snooping globally.
```

```
<SwitchC> system-view
[SwitchC] igmp-snooping
[SwitchC-igmp-snooping] quit
```

```
# Create VLAN 100, assign GigabitEthernet 1/0/1 through GigabitEthernet 1/0/5 to this VLAN,
and enable IGMP snooping in the VLAN.
```

```
[SwitchC] vlan 100
[SwitchC-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/5
[SwitchC-vlan100] igmp-snooping enable
[SwitchC-vlan100] quit
```

```
# Configure GigabitEthernet 1/0/3 and GigabitEthernet 1/0/5 as static member ports for
multicast group 224.1.1.1.
```

```
[SwitchC] interface GigabitEthernet 1/0/3
[SwitchC-GigabitEthernet1/0/3] igmp-snooping static-group 224.1.1.1 vlan 100
[SwitchC-GigabitEthernet1/0/3] quit
[SwitchC] interface GigabitEthernet 1/0/5
[SwitchC-GigabitEthernet1/0/5] igmp-snooping static-group 224.1.1.1 vlan 100
[SwitchC-GigabitEthernet1/0/5] quit
```

6. Verify the configuration:

```
# Display detailed IGMP snooping group information in VLAN 100 on Switch A.
```

```
[SwitchA] display igmp-snooping group vlan 100 verbose
```

```
Total 1 IP Group(s).
```

```
Total 1 IP Source(s).
```

```
Total 1 MAC Group(s).
```

```
Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port
```

```
Subvlan flags: R-Real VLAN, C-Copy VLAN
```

```
Vlan(id):100.
```

```
Total 1 IP Group(s).
```

```
Total 1 IP Source(s).
```

```
Total 1 MAC Group(s).
```



```

Router port(s):total 2 port.
    GE1/0/1                (D) ( 00:01:30 )
    GE1/0/3                (S)
IP group(s):the following ip group(s) match to one mac group.
IP group address:224.1.1.1
(0.0.0.0, 224.1.1.1):
Attribute:    Host Port
Host port(s):total 1 port.
    GE1/0/2                (D) ( 00:03:23 )
MAC group(s):
MAC group address:0100-5e01-0101
Host port(s):total 1 port.
    GE1/0/2

```

The output shows that GigabitEthernet 1/0/3 of Switch A has become a static router port.

Display detailed IGMP snooping group information in VLAN 100 on Switch C.

```
[SwitchC] display igmp-snooping group vlan 100 verbose
```

```

Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).

```

```

Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port
Subvlan flags: R-Real VLAN, C-Copy VLAN

```

```
Vlan(id):100.
```

```

Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).
Router port(s):total 1 port.
    GE1/0/2                (D) ( 00:01:23 )
IP group(s):the following ip group(s) match to one mac group.
IP group address:224.1.1.1
(0.0.0.0, 224.1.1.1):
Attribute:    Host Port
Host port(s):total 2 port.
    GE1/0/3                (S)
    GE1/0/5                (S)
MAC group(s):
MAC group address:0100-5e01-0101
Host port(s):total 2 port.
    GE1/0/3
    GE1/0/5

```

The output shows that GigabitEthernet 1/0/3 and GigabitEthernet 1/0/5 on Switch C have become static member ports for multicast group 224.1.1.1.

IGMP snooping querier configuration example

Network requirements

As shown in [Figure 16](#), in a Layer 2-only network environment, two multicast sources Source 1 and Source 2 send multicast data to multicast groups 224.1.1.1 and 225.1.1.1 respectively, Host A and Host C

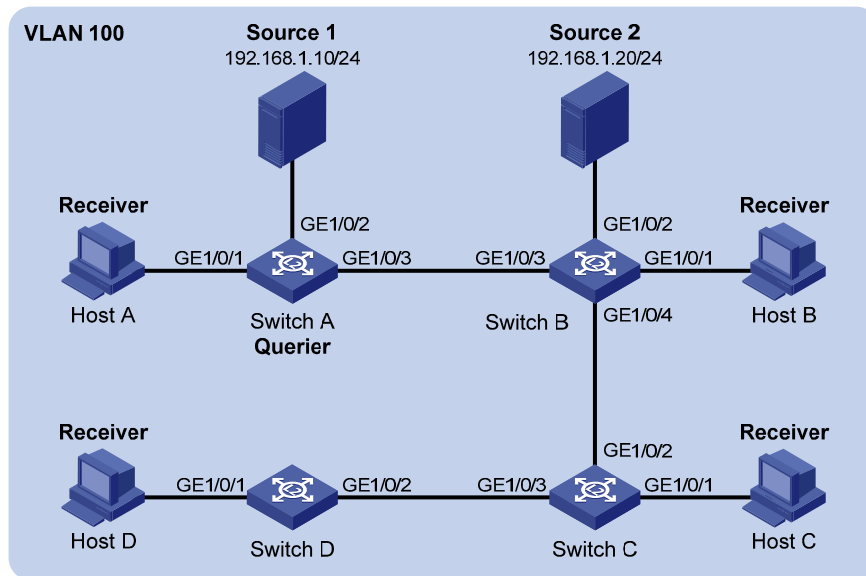
are receivers of multicast group 224.1.1.1, and Host B and Host D are receivers of multicast group 225.1.1.1.

All the receivers run IGMPv2, and all the switches run IGMPv2 snooping. Switch A, which is close to the multicast sources, is chosen as the IGMP snooping querier.

To prevent flooding of unknown multicast traffic within the VLAN, be sure to configure all the switches to drop unknown multicast data packets.

Because a switch does not enlist a port that has heard an IGMP query with a source IP address of 0.0.0.0 (default) as a dynamic router port, configure a non-all-zero IP address as the source IP address of IGMP queries to ensure normal creation of Layer 2 multicast forwarding entries.

Figure 16 Network diagram



Configuration procedure

1. Configure switch A:

Enable IGMP snooping globally.

```
<SwitchA> system-view
[SwitchA] igmp-snooping
[SwitchA-igmp-snooping] quit
```

Create VLAN 100 and assign GigabitEthernet 1/0/1 through GigabitEthernet 1/0/3 to the VLAN.

```
[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/3
```

Enable IGMP snooping and the function of dropping unknown multicast traffic in VLAN 100.

```
[SwitchA-vlan100] igmp-snooping enable
[SwitchA-vlan100] igmp-snooping drop-unknown
```

Enable the IGMP snooping querier function in VLAN 100

```
[SwitchA-vlan100] igmp-snooping querier
```

Set the source IP address of IGMP general queries and group-specific queries to 192.168.1.1 in VLAN 100.

```
[SwitchA-vlan100] igmp-snooping general-query source-ip 192.168.1.1
[SwitchA-vlan100] quit
```

2. Configure Switch B:

Enable IGMP snooping globally.

```
<SwitchB> system-view
[SwitchB] igmp-snooping
[SwitchB-igmp-snooping] quit
```

Create VLAN 100, and assign GigabitEthernet 1/0/1 through GigabitEthernet 1/0/4 to the VLAN.

```
[SwitchB] vlan 100
[SwitchB-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/4
```

Enable IGMP snooping and the function of dropping unknown multicast traffic in VLAN 100.

```
[SwitchB-vlan100] igmp-snooping enable
[SwitchB-vlan100] igmp-snooping drop-unknown
[SwitchB-vlan100] quit
```

Configurations on Switch C and Switch D are similar to the configuration on Switch B.

3. Verify the configuration:

After the IGMP snooping querier starts to work, all the switches but the querier can receive IGMP general queries. By using the **display igmp-snooping statistics** command, you can view the statistics information about the IGMP messages received. For example:

Display IGMP message statistics on Switch B.

```
[SwitchB] display igmp-snooping statistics
Received IGMP general queries:3.
Received IGMPv1 reports:0.
Received IGMPv2 reports:12.
Received IGMP leaves:0.
Received IGMPv2 specific queries:0.
Sent IGMPv2 specific queries:0.
Received IGMPv3 reports:0.
Received IGMPv3 reports with right and wrong records:0.
Received IGMPv3 specific queries:0.
Received IGMPv3 specific sg queries:0.
Sent IGMPv3 specific queries:0.
Sent IGMPv3 specific sg queries:0.
Received error IGMP messages:0.
```

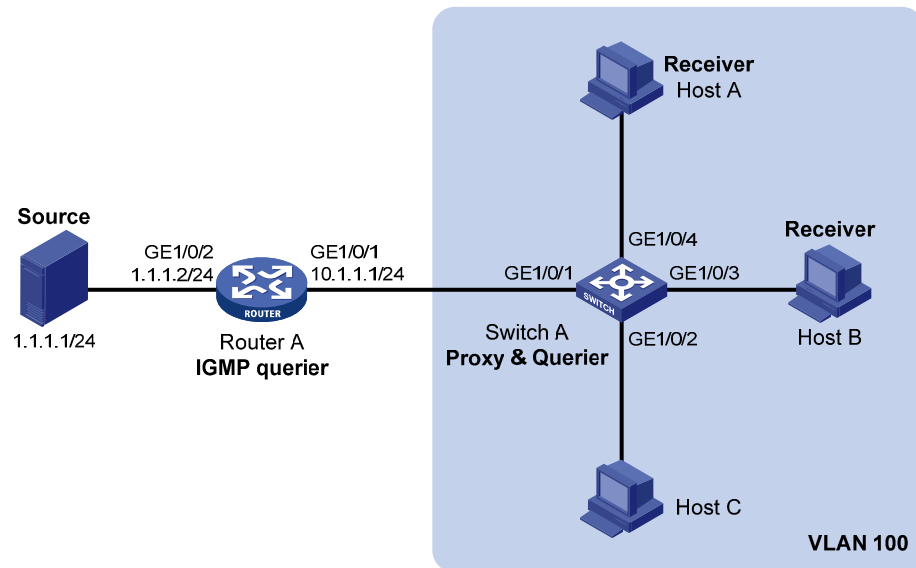
IGMP snooping proxying configuration example

Network requirements

As shown in [Figure 17](#), Router A runs IGMPv2 and Switch A runs IGMPv2 snooping. Router A acts as the IGMP querier.

Configure IGMP snooping proxying on Switch A, enabling the switch to forward IGMP reports and leave messages on behalf of attached hosts and to respond to IGMP queries from Router A and forward the queries to the hosts on behalf of Router A.

Figure 17 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces:
Configure an IP address and subnet mask for each interface as per Figure 17. (Details not shown.)
2. Configure Router A:
Enable IP multicast routing, enable PIM-DM on each interface, and enable IGMP on GigabitEthernet 1/0/1.

```
<RouterA> system-view
[RouterA] multicast routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] igmp enable
[RouterA-GigabitEthernet1/0/1] pim dm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] pim dm
[RouterA-GigabitEthernet1/0/2] quit
```
3. Configure Switch A:
Enable IGMP snooping globally.

```
<SwitchA> system-view
[SwitchA] igmp-snooping
[SwitchA-igmp-snooping] quit
```


Create VLAN 100, assign ports GigabitEthernet 1/0/1 through GigabitEthernet 1/0/4 to this VLAN, and enable IGMP snooping and IGMP snooping proxying in the VLAN.

```
[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/4
[SwitchA-vlan100] igmp-snooping enable
[SwitchA-vlan100] igmp-snooping proxying enable
[SwitchA-vlan100] quit
```
4. Verify the configuration:

After the configuration is completed, Host A and Host B send IGMP join messages for group 224.1.1.1. Receiving the messages, Switch A sends a join message for the group out of port GigabitEthernet 1/0/1 (a router port) to Router A.

Use the **display igmp-snooping group** command and the **display igmp group** command to display information about IGMP snooping groups and IGMP multicast groups. For example:

Display information about IGMP snooping groups on Switch A.

```
[SwitchA] display igmp-snooping group
Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).

Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port
Subvlan flags: R-Real VLAN, C-Copy VLAN
Vlan(id):100.
Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).
Router port(s):total 1 port.
    GE1/0/1                (D) ( 00:01:23 )
IP group(s):the following ip group(s) match to one mac group.
IP group address:224.1.1.1
(0.0.0.0, 224.1.1.1):
Host port(s):total 2 port.
    GE1/0/3                (D)
    GE1/0/4                (D)
MAC group(s):
MAC group address:0100-5e01-0101
Host port(s):total 2 port.
    GE1/0/3
    GE1/0/4
```

Display information about IGMP multicast groups on Router A.

```
[RouterA] display igmp group
Total 1 IGMP Group(s).
Interface group report information of VPN-Instance: public net
GigabitEthernet1/0/1(10.1.1.1):
Total 1 IGMP Group reported
Group Address      Last Reporter      Uptime      Expires
224.1.1.1          0.0.0.0            00:00:06    00:02:04
```

When Host A leaves the multicast group, it sends an IGMP leave message to Switch A. Receiving the message, Switch A removes port GigabitEthernet 1/0/4 from the member port list of the forwarding entry for the group; however, it does not remove the group or forward the leave message to Router A because Host B is still in the group. Use the **display igmp-snooping group** command to display information about IGMP snooping groups. For example:

Display information about IGMP snooping groups on Switch A.

```
[SwitchA] display igmp-snooping group
Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).
```

```

Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port
Subvlan flags: R-Real VLAN, C-Copy VLAN
Vlan(id):100.
Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).
Router port(s):total 1 port.
    GE1/0/1                (D) ( 00:01:23 )
IP group(s):the following ip group(s) match to one mac group.
IP group address:224.1.1.1
(0.0.0.0, 224.1.1.1):
Host port(s):total 1 port.
    GE1/0/3                (D)
MAC group(s):
MAC group address:0100-5e01-0101
Host port(s):total 1 port.
    GE1/0/3

```

Multicast source and user control policy configuration example

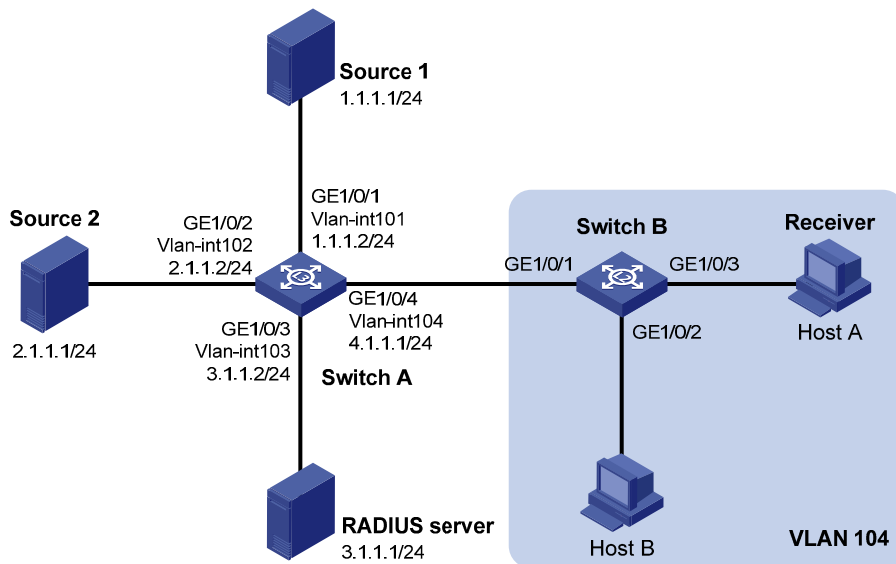
Network requirements

As shown in [Figure 18](#), Switch A is a Layer-3 switch. Switch A runs IGMPv2 and Switch B runs IGMPv2 snooping. Multicast sources and hosts run 802.1X client.

A multicast source control policy is configured on Switch A to block multicast flows from Source 2 to 224.1.1.1.

A multicast user control policy is configured on Switch B so that Host A can join or leave only multicast group 224.1.1.1.

Figure 18 Network diagram



Configuration procedures

1. Configure IP addresses for interfaces:

Configure an IP address and subnet mask for each interface as per [Figure 18](#). (Details not shown.)

2. Configure Switch A:

Create VLAN 101 through VLAN 104 and assign GigabitEthernet 1/0/1 through GigabitEthernet 1/0/4 to the four VLANs respectively.

```
<SwitchA> system-view
[SwitchA] vlan 101
[SwitchA-vlan101] port gigabitethernet 1/0/1
[SwitchA-vlan101] quit
[SwitchA] vlan 102
[SwitchA-vlan102] port gigabitethernet 1/0/2
[SwitchA-vlan102] quit
[SwitchA] vlan 103
[SwitchA-vlan103] port gigabitethernet 1/0/3
[SwitchA-vlan103] quit
[SwitchA] vlan 104
[SwitchA-vlan104] port gigabitethernet 1/0/4
[SwitchA-vlan104] quit
```

Enable IP multicast routing. Enable PIM-DM on VLAN-interface 101, VLAN-interface 102 and VLAN-interface 104, and enable IGMP on VLAN-interface 104.

```
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim dm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 102
[SwitchA-Vlan-interface102] pim dm
[SwitchA-Vlan-interface102] quit
[SwitchA] interface vlan-interface 104
[SwitchA-Vlan-interface104] pim dm
[SwitchA-Vlan-interface104] igmp enable
[SwitchA-Vlan-interface104] quit
```

Create QoS policy **policy1** to block multicast flows from Source 2 to 224.1.1.1.

```
[SwitchA] acl number 3001
[SwitchA-acl-adv-3001] rule permit udp source 2.1.1.1 0 destination 224.1.1.1 0
[SwitchA-acl-adv-3001] quit [SwitchA] traffic classifier classifier1
[SwitchA-classifier-classifier1] if-match acl 3001
[SwitchA-classifier-classifier1] quit
[SwitchA] traffic behavior behavior1
[SwitchA-behavior-behavior1] filter deny
[SwitchA-behavior-behavior1] quit
[SwitchA] qos policy policy1
[SwitchA-qospolicy-policy1] classifier classifier1 behavior behavior1
[SwitchA-qospolicy-policy1] quit
```

Create user profile **profile1**, apply QoS policy **policy1** to the inbound direction in user profile view, and enable the user profile.

```
[SwitchA] user-profile profile1
[SwitchA-user-profile-profile1] qos apply policy policy1 inbound
[SwitchA-user-profile-profile1] quit
```

```
[SwitchA] user-profile profile1 enable
```

Create RADIUS scheme **scheme1**; set the service type for the RADIUS server to **extended**; specify the IP addresses of the primary authentication/authorization server and accounting server as 3.1.1.1; set the shared keys to 123321; specify that no domain name is carried in a username sent to the RADIUS server.

```
[SwitchA] radius scheme scheme1
[SwitchA-radius-scheme1] server-type extended
[SwitchA-radius-scheme1] primary authentication 3.1.1.1
[SwitchA-radius-scheme1] key authentication 123321
[SwitchA-radius-scheme1] primary accounting 3.1.1.1
[SwitchA-radius-scheme1] key accounting 123321
[SwitchA-radius-scheme1] user-name-format without-domain
[SwitchA-radius-scheme1] quit
```

Create ISP domain **domain1**; reference **scheme1** for the authentication, authorization, and accounting of LAN users; specify **domain1** as the default ISP domain.

```
[SwitchA] domain domain1
[SwitchA-isp-domain1] authentication lan-access radius-scheme scheme1
[SwitchA-isp-domain1] authorization lan-access radius-scheme scheme1
[SwitchA-isp-domain1] accounting lan-access radius-scheme scheme1
[SwitchA-isp-domain1] quit
[SwitchA] domain default enable domain1
```

Globally enable 802.1X and then enable it on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 respectively.

```
[SwitchA] dot1x
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] dot1x
[SwitchA-GigabitEthernet1/0/1] quit
[SwitchA] interface gigabitethernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] dot1x
[SwitchA-GigabitEthernet1/0/2] quit
```

3. Configure Switch B:

Globally enable IGMP snooping.

```
<SwitchB> system-view
[SwitchB] igmp-snooping
[SwitchB-igmp-snooping] quit
```

Create VLAN 104, assign GigabitEthernet 1/0/1 through GigabitEthernet 1/0/3 to this VLAN, and enable IGMP snooping in this VLAN.

```
[SwitchB] vlan 104
[SwitchB-vlan104] port gigabitethernet 1/0/1 to gigabitethernet 1/0/3
[SwitchB-vlan104] igmp-snooping enable
[SwitchB-vlan104] quit
```

Create a user profile **profile2** to allow users to join or leave only one multicast group, 224.1.1.1. Then, enable the user profile.

```
[SwitchB] acl number 2001
[SwitchB-acl-basic-2001] rule permit source 224.1.1.1 0
[SwitchB-acl-basic-2001] quit
[SwitchB] user-profile profile2
```



```
[SwitchB-user-profile-profile2] igmp-snooping access-policy 2001
[SwitchB-user-profile-profile2] quit
[SwitchB] user-profile profile2 enable
```

Create a RADIUS scheme **scheme2**; set the service type for the RADIUS server to **extended**; specify the IP addresses of the primary authentication/authorization server and accounting server as 3.1.1.1; set the shared keys to 321123; specify that a username sent to the RADIUS server carry no domain name.

```
[SwitchB] radius scheme scheme2
[SwitchB-radius-scheme2] server-type extended
[SwitchB-radius-scheme2] primary authentication 3.1.1.1
[SwitchB-radius-scheme2] key authentication 321123
[SwitchB-radius-scheme2] primary accounting 3.1.1.1
[SwitchB-radius-scheme2] key accounting 321123
[SwitchB-radius-scheme2] user-name-format without-domain
[SwitchB-radius-scheme2] quit
```

Create an ISP domain **domain2**; reference **scheme2** for the authentication, authorization, and accounting of LAN users; specify **domain2** as the default ISP domain.

```
[SwitchB] domain domain2
[SwitchB-isp-domian2] authentication lan-access radius-scheme scheme2
[SwitchB-isp-domian2] authorization lan-access radius-scheme scheme2
[SwitchB-isp-domian2] accounting lan-access radius-scheme scheme2
[SwitchB-isp-domian2] quit
[SwitchB] domain default enable domain2
```

Globally enable 802.1X and then enable it on GigabitEthernet 1/0/2 and GigabitEthernet 1/0/3 respectively.

```
[SwitchB] dot1x
[SwitchB] interface gigabitethernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] dot1x
[SwitchB-GigabitEthernet1/0/2] quit
[SwitchB] interface gigabitethernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] dot1x
[SwitchB-GigabitEthernet1/0/3] quit
```

4. Configure the RADIUS server:

On the RADIUS server, configure the parameters related to Switch A and Switch B. For more information, see the configuration guide of the RADIUS server.

5. Verify the configuration:

After the configurations, the two multicast sources and hosts initiate 802.1X authentication. After passing authentication, Source 1 sends multicast flows to 224.1.1.1 and Source 2 sends multicast flows to 224.1.1.2; Host A sends messages to join multicast groups 224.1.1.1 and 224.1.1.2. Use the **display igmp-snooping group** command to display information about IGMP snooping groups. For example:

Display information about IGMP snooping groups in VLAN 104 on Switch B.

```
[SwitchB] display igmp-snooping group vlan 104 verbose
Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).
```

```

Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port
Subvlan flags: R-Real VLAN, C-Copy VLAN
Vlan(id):104.
Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).
Router port(s):total 1 port.
    GE1/0/1                (D) ( 00:01:30 )
IP group(s):the following ip group(s) match to one mac group.
  IP group address:224.1.1.1
    (0.0.0.0, 224.1.1.1):
      Attribute:    Host Port
      Host port(s):total 1 port.
        GE1/0/3                (D) ( 00:04:10 )
MAC group(s):
  MAC group address:0100-5e01-0101
  Host port(s):total 1 port.
    GE1/0/3

```

The output shows that GigabitEthernet 1/0/3 on Switch B has joined 224.1.1.1 but not 224.1.1.2.

Assume that Source 2 starts sending multicast traffic to 224.1.1.1. Use the **display multicast forwarding-table** to display the multicast forwarding table information.

Display information about 224.1.1.1 in the multicast forwarding table on Switch A.

```

[SwitchA] display multicast forwarding-table 224.1.1.1
Multicast Forwarding Table of VPN-Instance: public net

```

```
Total 1 entry
```

```
Total 1 entry matched
```

```

00001. (1.1.1.1, 224.1.1.1)
  MID: 0, Flags: 0x0:0
  Uptime: 00:08:32, Timeout in: 00:03:26
  Incoming interface: Vlan-interface101
  List of 1 outgoing interfaces:
    1: Vlan-interface104
  Matched 19648 packets(20512512 bytes), Wrong If 0 packets
  Forwarded 19648 packets(20512512 bytes)

```

The output shows that Switch A maintains a multicast forwarding entry for multicast packets from Source 1 to 224.1.1.1. No forwarding entry exists for packets from Source 2 to 224.1.1.1, which indicates that multicast packets from Source 2 are blocked.

Troubleshooting IGMP snooping

Layer 2 multicast forwarding cannot function

Symptom

Layer 2 multicast forwarding cannot function.

Analysis

IGMP snooping is not enabled.

Solution

1. Use the **display current-configuration** command to check the running status of IGMP snooping.
2. If IGMP snooping is not enabled, use the **igmp-snooping** command to enable IGMP snooping globally, and then use the **igmp-snooping enable** command to enable IGMP snooping in VLAN view.
3. If IGMP snooping is disabled only for the corresponding VLAN, use the **igmp-snooping enable** command in VLAN view to enable IGMP snooping in the corresponding VLAN.

Configured multicast group policy fails to take effect

Symptom

Although a multicast group policy has been configured to allow hosts to join specific multicast groups, the hosts can still receive multicast data addressed to other multicast groups.

Analysis

- The ACL rule is incorrectly configured.
- The multicast group policy is not correctly applied.
- The function of dropping unknown multicast data is not enabled, so unknown multicast data is flooded.

Solution

1. Use the **display acl** command to check the configured ACL rule. Make sure that the ACL rule conforms to the multicast group policy to be implemented.
2. Use the **display this** command in IGMP-snooping view or in the corresponding interface view to verify that the correct multicast group policy has been applied. If not, use the **group-policy** or **igmp-snooping group-policy** command to apply the correct multicast group policy.
3. Use the **display current-configuration** command to verify that the function of dropping unknown multicast data is enabled. If not, use the **drop-unknown** or **igmp-snooping drop-unknown** command to enable the function of dropping unknown multicast data.

Appendix

Processing of multicast protocol messages

With Layer 3 multicast routing enabled, an IGMP snooping-enabled switch processes multicast protocol messages differently under different conditions, as follows:

1. If only IGMP is enabled on the switch, or if both IGMP and PIM are enabled on the switch, the switch does the following:
 - Maintains dynamic member ports or dynamic router ports according to IGMP packets
 - Maintains dynamic router ports according to PIM hello packets
2. If only PIM is enabled on the switch, the following occur:
 - The switch broadcasts IGMP messages as unknown messages in the VLAN.
 - After receiving a PIM hello message, the switch maintains the corresponding dynamic router port.
3. If IGMP is disabled on the switch, one of the following occurs:
 - If PIM is disabled, the switch deletes all its dynamic member ports and dynamic router ports.
 - If PIM is enabled, the switch deletes only its dynamic member ports but not its dynamic router ports.

NOTE:

On a switch with Layer-3 multicast routing enabled, use the **display igmp group port-info** command to display Layer-2 port information. For more information about this command, see *IP Multicast Command Reference*.

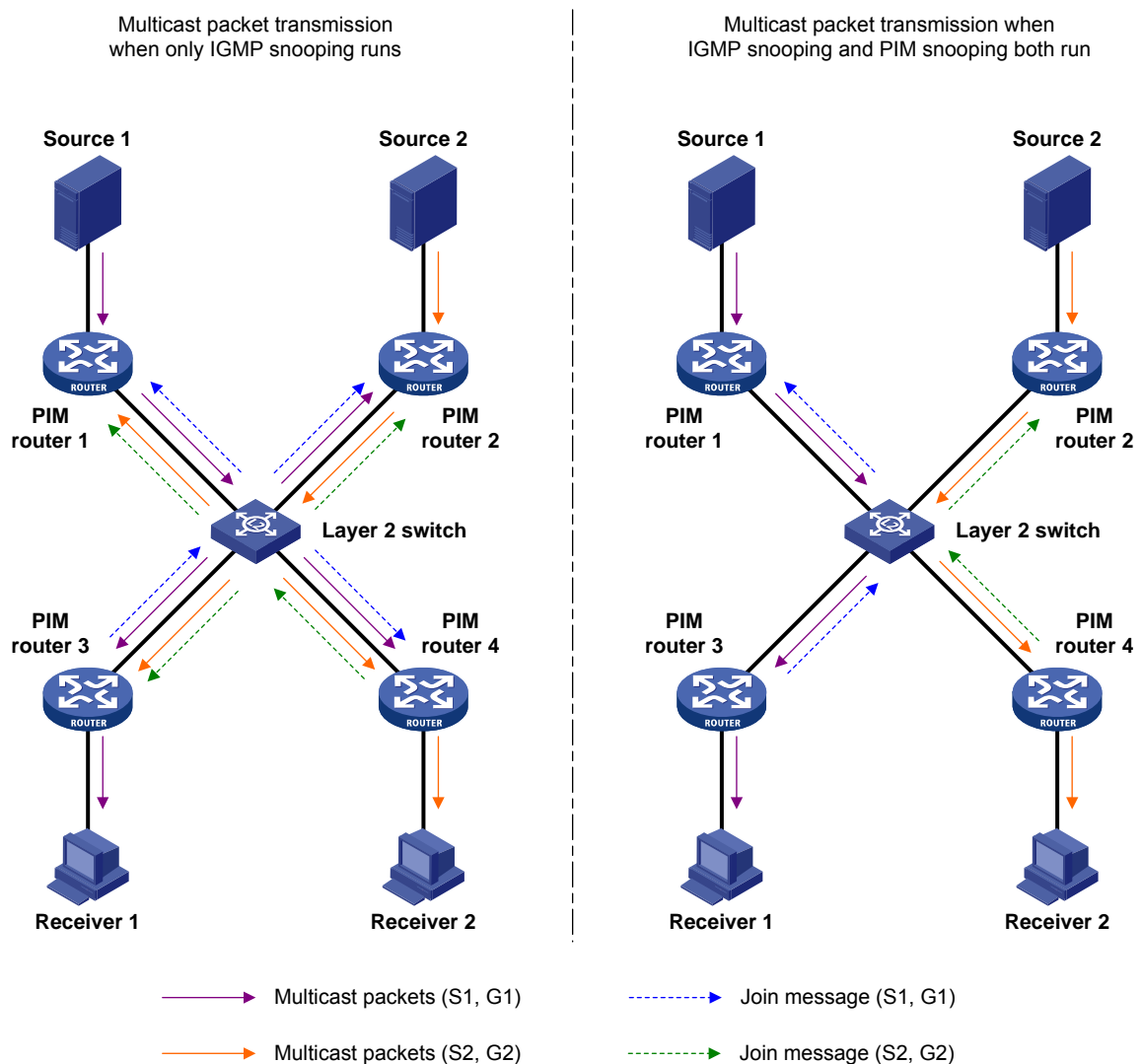
4. If PIM is disabled on the switch, one of the following occurs:
 - If IGMP is disabled, the switch deletes all its dynamic router ports.
 - If IGMP is enabled, the switch maintains all its dynamic member ports and dynamic router ports.

Configuring PIM snooping

Overview

Protocol Independent Multicast (PIM) snooping runs on Layer 2 devices. It determines which ports are interested in multicast data by analyzing the received PIM messages, and adds the ports to a multicast forwarding entry to make sure that multicast data can be forwarded to only the ports that are interested in the data.

Figure 19 Multicast packet transmission without or with PIM snooping



As shown in Figure 19, Source 1 sends multicast data to multicast group G1, and Source 2 sends multicast data to multicast group G2. Receiver 1 belongs to G1, and Receiver 2 belongs to G2. The Layer 2 switch's interfaces that connect to the PIM-capable routers are in the same VLAN.

- When the Layer 2 switch runs only IGMP snooping, it maintains the router ports according to the received PIM hello messages that PIM-capable routers send, broadcasts all other types of received

PIM messages in the VLAN, and forwards all multicast data to all router ports in the VLAN. Each PIM-capable router in the VLAN, whether interested in the multicast data or not, can receive all multicast data and all PIM messages except PIM hello messages.

- When the Layer 2 switch runs both IGMP snooping and PIM snooping, it determines whether PIM-capable routers are interested in the multicast data addressed to a multicast group according to PIM messages received from the routers, and adds only the ports for connecting the routers that are interested in the data to a multicast forwarding entry. Then, the Layer 2 switch forwards PIM messages and multicast data to only the routers that are interested in the data, saving network bandwidth.

For more information about IGMP snooping and the router port, see "[Configuring IGMP snooping](#) ."

For more information about PIM, see "[Configuring PIM \(available only on the HP 5500 EI\)](#)"

Configuring PIM snooping

Configuration guidelines

Before configuring PIM snooping for a VLAN, be sure to enable IGMP snooping globally and specifically for the VLAN.

After you enable PIM snooping in a VLAN, PIM snooping works only on the member interfaces of the VLAN.

PIM snooping does not work in the sub-VLANs of a multicast VLAN. For more information about multicast VLAN, see "[Configuring multicast VLANs](#)."

In a network with PIM snooping enabled switches, configure the size of each join/prune message no more than the path maximum transmission unit (MTU) on the PIM-enabled edge router on the receiver side. For more information about the join/prune messages, see "[Configuring PIM \(available only on the HP 5500 EI\)](#)."

Configuration procedure

To configure PIM snooping:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IGMP snooping globally and enter IGMP-snooping view.	igmp-snooping	Disabled by default
3. Return to system view.	quit	N/A
4. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
5. Enable IGMP snooping in the VLAN.	igmp-snooping enable	Disabled by default
6. Enable PIM snooping in the VLAN.	pim-snooping enable	Disabled by default

For more information about the **igmp-snooping** and **igmp-snooping enable** commands, see *IP Multicast Command Reference*.

Displaying and maintaining PIM snooping

Task	Command	Remarks
Display PIM snooping neighbor information.	display pim-snooping neighbor [vlan <i>vlan-id</i>] [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display PIM snooping routing entries.	display pim-snooping routing-table [vlan <i>vlan-id</i>] [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the statistics information of PIM messages learned by PIM snooping.	display pim-snooping statistics [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear the statistics information of PIM messages learned by PIM snooping.	reset pim-snooping statistics	Available in user view

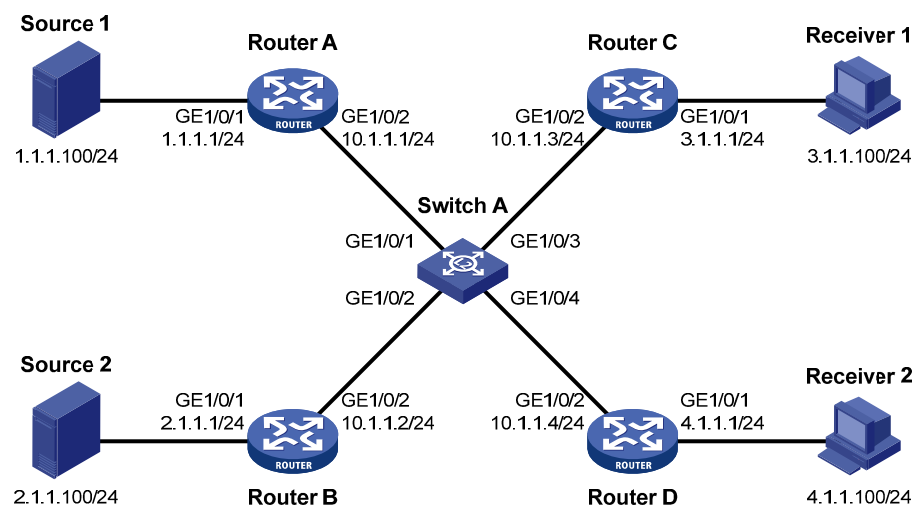
PIM snooping configuration example

Network requirements

As shown in Figure 20, Source 1 sends multicast data to multicast group 224.1.1.1, and Source 2 sends multicast data to multicast group 225.1.1.1. Receiver 1 belongs to multicast group 224.1.1.1, and Receiver 2 belongs to multicast group 225.1.1.1. Router C and Router D run IGMP on their interface GigabitEthernet 1/0/1. Router A, Router B, Router C, and Router D run PIM-SM, and interface GigabitEthernet 1/0/2 on Router A acts as a C-BSR and C-RP.

Configure IGMP snooping and PIM snooping on Switch A so that Switch A forwards PIM messages and multicast data to only the routers that are interested in the multicast data.

Figure 20 Network diagram



Configuration procedure

1. Assign IP addresses:

Configure an IP address and subnet mask for each interface according to [Figure 20](#). (Details not shown.)

2. Configure Router A:

Enable IP multicast routing, enable PIM-SM on each interface, and configure interface GigabitEthernet 1/0/2 as a C-BSR and C-RP.

```
<RouterA> system-view
[RouterA] multicast routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] pim sm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] pim sm
[RouterA-GigabitEthernet1/0/2] quit
[RouterA] pim
[RouterA-pim] c-bsr gigabitethernet 1/0/2
[RouterA-pim] c-rp gigabitethernet 1/0/2
```

3. Configure Router B:

Enable IP multicast routing, and enable PIM-SM on each interface.

```
<RouterB> system-view
[RouterB] multicast routing-enable
[RouterB] interface gigabitethernet 1/0/1
[RouterB-GigabitEthernet1/0/1] pim sm
[RouterB-GigabitEthernet1/0/1] quit
[RouterB] interface gigabitethernet 1/0/2
[RouterB-GigabitEthernet1/0/2] pim sm
```

4. Configure Router C:

Enable IP multicast routing, enable PIM-SM on each interface, and enable IGMP on GigabitEthernet 1/0/1.

```
<RouterC> system-view
[RouterC] multicast routing-enable
[RouterC] interface gigabitethernet 1/0/1
[RouterC-GigabitEthernet1/0/1] pim sm
[RouterC-GigabitEthernet1/0/1] igmp enable
[RouterC-GigabitEthernet1/0/1] quit
[RouterC] interface gigabitethernet 1/0/2
[RouterC-GigabitEthernet1/0/2] pim sm
```

5. Configure Router D:

The configuration on Router D is similar to that on Router C. (Details not shown.)

6. Configure Switch A:

Enable IGMP snooping globally.

```
<SwitchA> system-view
[SwitchA] igmp-snooping
[SwitchA-igmp-snooping] quit
```


Create VLAN 100, assign GigabitEthernet 1/0/1 through GigabitEthernet 1/0/4 to this VLAN, and enable IGMP snooping and PIM snooping in the VLAN.

```
[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/4
[SwitchA-vlan100] igmp-snooping enable
[SwitchA-vlan100] pim-snooping enable
[SwitchA-vlan100] quit
```

7. Verify the configuration:

On Switch A, display the PIM snooping neighbor information of VLAN 100.

```
[SwitchA] display pim-snooping neighbor vlan 100
Total number of neighbors: 4
```

```
VLAN ID: 100
Total number of neighbors: 4
Neighbor      Port          Expires      Option Flags
10.1.1.1      GE1/0/1       02:02:23    LAN Prune Delay
10.1.1.2      GE1/0/2       03:00:05    LAN Prune Delay
10.1.1.3      GE1/0/3       02:22:13    LAN Prune Delay
10.1.1.4      GE1/0/4       03:07:22    LAN Prune Delay
```

The output shows that Router A, Router B, Router C, and Router D are PIM snooping neighbors.

On Switch A, display the PIM snooping routing information of VLAN 100.

```
[SwitchA] display pim-snooping routing-table vlan 100 slot 1
Total 2 entry(ies)
FSM Flag: NI-no info, J-join, PP-prune pending
```

```
VLAN ID: 100
Total 2 entry(ies)
(*, 224.1.1.1)
  Upstream neighbor: 10.1.1.1
  Upstream port: GE1/0/1
  Total number of downstream ports: 1
    1: GE1/0/3
      Expires: 00:03:01, FSM: J
(*, 225.1.1.1)
  Upstream neighbor: 10.1.1.1
  Upstream port: GE1/0/2
  Total number of downstream ports: 1
    1: GE1/0/4
      Expires: 00:01:05, FSM: J
```

The output shows that Switch A will forward the multicast data intended for multicast group 224.1.1.1 to only Router C, and forward the multicast data intended for multicast group 225.1.1.1 to only Router D.

Troubleshooting PIM snooping

PIM snooping does not work

Symptom

PIM snooping does not work on the switch.

Analysis

IGMP snooping or PIM snooping is not enabled on the switch.

Solution

1. Use the **display current-configuration** command to check the status of IGMP snooping and PIM snooping.
2. If IGMP snooping is not enabled, enter system view and use the **igmp-snooping** command to enable IGMP snooping globally. Then, enter VLAN view and use the **igmp-snooping enable** and **pim-snooping enable** commands to enable IGMP snooping and PIM snooping in the VLAN.
3. If PIM snooping is not enabled, enter VLAN view and use the **pim-snooping enable** command to enable PIM snooping in the VLAN.

Some downstream PIM-capable routers cannot receive multicast data

Symptom

In a network with fragmented join/prune messages, some downstream PIM-capable routers cannot receive multicast data.

Analysis

PIM snooping cannot reassemble messages, and it cannot maintain the status of downstream routers that the join/prune message fragments carry. To ensure the normal operation of the system, PIM snooping must broadcast join/prune message fragments in the VLAN. However, if the VLAN has a PIM-capable router that has the join suppression function enabled, the broadcast join/prune message fragments might suppress the join messages of other PIM-capable routers in the VLAN. As a result, some PIM-capable routers cannot receive the multicast data destined for a specific multicast group because their join messages are suppressed. To solve this problem, disable the join suppression function on all PIM-capable routers that connect to the PIM snooping-capable switch in the VLAN.

Solution

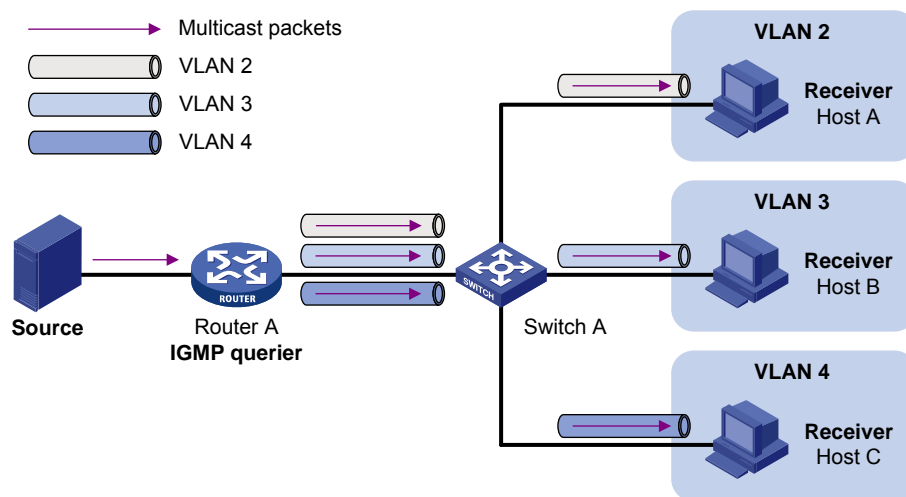
1. Use the **pim hello-option neighbor-tracking** command to enable the neighbor tracking function on the interfaces of PIM routers that connect to the PIM snooping-capable switch.
2. If a PIM-capable router cannot be enabled with the neighbor tracking function, you have to disable PIM snooping on the switch.

Configuring multicast VLANs

Overview

In the traditional multicast programs-on-demand mode shown in [Figure 21](#), when hosts (Host A, Host B and Host C) that belong to different VLANs require multicast programs-on-demand service, the Layer 3 device, Router A, must forward a separate copy of the multicast traffic in each user VLAN to the Layer 2 device, Switch A. This results in not only waste of network bandwidth but also extra burden on the Layer 3 device.

Figure 21 Multicast transmission without multicast VLAN



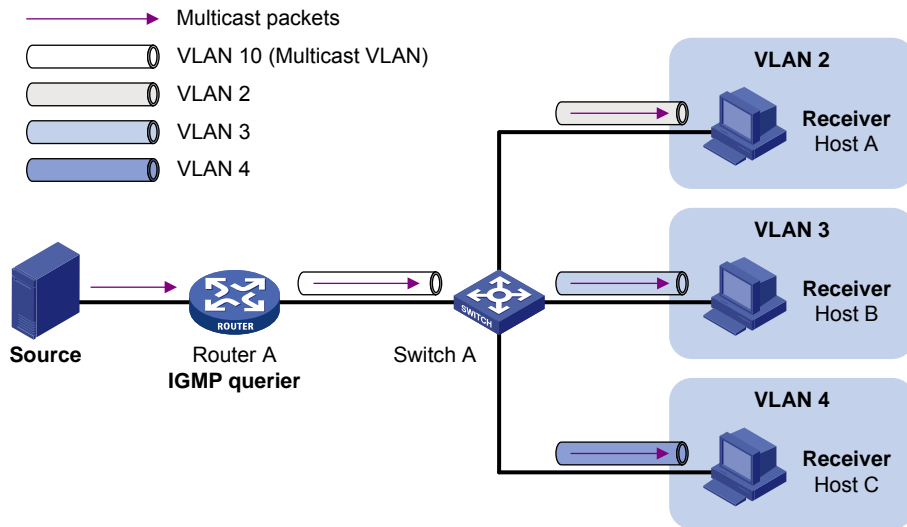
The multicast VLAN feature configured on the Layer 2 device is the solution to this issue. With the multicast VLAN feature, the Layer 3 device replicates the multicast traffic only in the multicast VLAN instead of making a separate copy of the multicast traffic in each user VLAN. This saves network bandwidth and lessens the burden on the Layer 3 device.

The multicast VLAN feature can be implemented in sub-VLAN-based multicast VLAN and port-based multicast VLAN.

Sub-VLAN-based multicast VLAN

As shown in [Figure 22](#), Host A, Host B, and Host C are in different user VLANs. On Switch A, configure VLAN 10 as a multicast VLAN, configure all the user VLANs as sub-VLANs of VLAN 10, and enable IGMP snooping in the multicast VLAN.

Figure 22 Sub-VLAN-based multicast VLAN

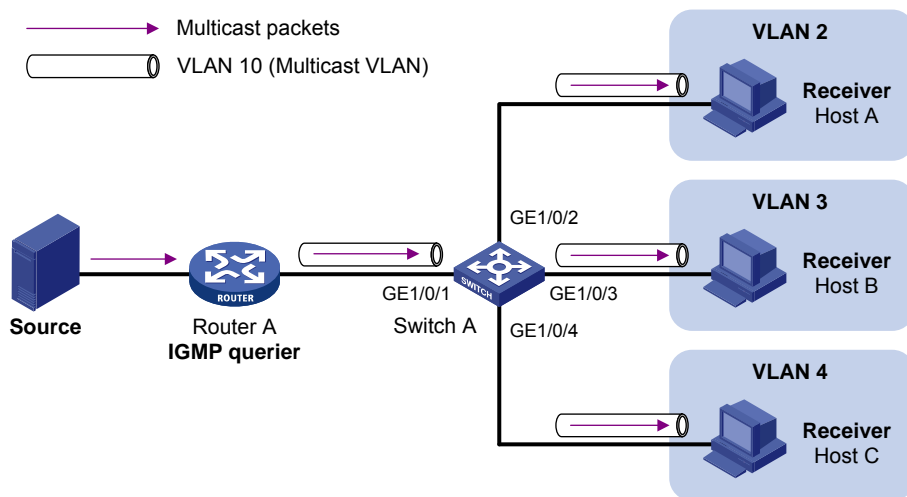


After the configuration, IGMP snooping manages router ports in the multicast VLAN and member ports in the sub-VLANs. When forwarding multicast data to Switch A, Router A sends only one copy of multicast data to Switch A in the multicast VLAN, and Switch A distributes the data to the multicast VLAN's sub-VLANs that contain receivers.

Port-based multicast VLAN

As shown in Figure 23, Host A, Host B, and Host C are in different user VLANs. All the user ports (ports with attached hosts) on Switch A are hybrid ports. On Switch A, configure VLAN 10 as a multicast VLAN, assign all the user ports to VLAN 10, and enable IGMP snooping in the multicast VLAN and all the user VLANs.

Figure 23 Port-based multicast VLAN



After the configuration, if Switch A receives an IGMP message on a user port, it tags the message with the multicast VLAN ID and relays it to the IGMP querier, so that IGMP snooping can uniformly manage the router port and member ports in the multicast VLAN. When Router A forwards multicast data to Switch A, it sends only one copy of multicast data to Switch A in the multicast VLAN, and Switch A distributes the data to all the member ports in the multicast VLAN.

For more information about IGMP snooping, router ports, and member ports, see "[Configuring IGMP snooping](#)."

For more information about VLAN tags, see *Layer 2—LAN Switching Configuration Guide*.

Multicast VLAN configuration task list

Task	Remarks
Configuring a sub-VLAN-based multicast VLAN	Required
Configuring a port-based multicast VLAN	Configuring user port attributes Configuring multicast VLAN ports

NOTE:

If you have configured both sub-VLAN-based multicast VLAN and port-based multicast VLAN on a device, the port-based multicast VLAN configuration is given preference.

Configuring a sub-VLAN-based multicast VLAN

Configuration prerequisites

Before you configure sub-VLAN-based multicast VLAN, complete the following tasks:

- Create VLANs as required.
- Enable IGMP snooping in the VLAN to be configured as a multicast VLAN.

Configuration guidelines

- You cannot configure multicast VLAN on a device with IP multicast routing enabled.
- The VLAN to be configured as a multicast VLAN must exist.
- The VLANs to be configured as sub-VLANs of the multicast VLAN must exist and must not be multicast VLANs or sub-VLANs of any other multicast VLAN.
- The total number of sub-VLANs of a multicast VLAN must not exceed the maximum number the system can support.

Configuration procedure

In this approach, you configure a VLAN as a multicast VLAN and configure user VLANs as sub-VLANs of the multicast VLAN.

To configure a sub-VLAN-based multicast VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the specified VLAN as a multicast VLAN and enter multicast VLAN view.	multicast-vlan <i>vlan-id</i>	By default, a VLAN is not a multicast VLAN.

Step	Command	Remarks
3. Configure the specified VLANs as sub-VLANs of the multicast VLAN.	subvlan <i>vlan-list</i>	By default, a multicast VLAN has no sub-VLANs.

Configuring a port-based multicast VLAN

When you configure a port-based multicast VLAN, you must configure the attributes of each user port and then assign the ports to the multicast VLAN.

A user port can be configured as a multicast VLAN port only if it is an Ethernet port, or Layer 2 aggregate interface.

In Ethernet interface view or Layer 2 aggregate interface view, configurations that you make are effective on only the current port. In port group view, configurations that you make are effective on all ports in the current port group.

Configuration prerequisites

Before you configure a port-based multicast VLAN, complete the following tasks:

- Create VLANs as required.
- Enable IGMP snooping in the VLAN to be configured as a multicast VLAN.
- Enable IGMP snooping in all the user VLANs.

Configuring user port attributes

Configure the user ports as hybrid ports that permit packets of the specified user VLAN to pass, and configure the user VLAN to which the user ports belong as the default VLAN.

Configure the user ports to permit packets of the multicast VLAN to pass and untag the packets. Thus, after receiving multicast packets tagged with the multicast VLAN ID from the upstream device, the Layer 2 device untags the multicast packets and forwards them to its downstream device.

To configure user port attributes:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> • Enter interface view or port group view: interface <i>interface-type</i> <i>interface-number</i> • Enter interface view or port group view: port-group { manual <i>port-group-name</i> aggregation <i>agg-id</i> } 	Use either command.
3. Configure the user port link type as hybrid.	port link-type hybrid	Access by default

Step	Command	Remarks
4. Specify the user VLAN that comprises the current user ports as the default VLAN.	port hybrid pvid vlan <i>vlan-id</i>	VLAN 1 by default
5. Configure the current user ports to permit packets of the specified multicast VLANs to pass and untag the packets.	port hybrid vlan <i>vlan-id-list</i> untagged	By default, a hybrid port permits only packets of VLAN 1 to pass.

For more information about the **port link-type**, **port hybrid pvid vlan**, and **port hybrid vlan** commands, see *Layer 2—LAN Switching Command Reference*.

Configuring multicast VLAN ports

Configuration guidelines

In this approach, you configure a VLAN as a multicast VLAN and assign user ports to it. You can do this by either adding the user ports in the multicast VLAN or specifying the multicast VLAN on the user ports. These two methods provide the same result.

You cannot configure multicast VLAN on a device with multicast routing enabled.

The VLAN to be configured as a multicast VLAN must exist.

A port can belong to only one multicast VLAN.

Configuration procedure

To configure multicast VLAN ports in multicast VLAN view:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the specified VLAN as a multicast VLAN and enter multicast VLAN view.	multicast-vlan <i>vlan-id</i>	By default, a VLAN is not a multicast VLAN.
3. Assign ports to the multicast VLAN.	port <i>interface-list</i>	By default, a multicast VLAN has no ports.

To configure multicast VLAN ports in interface view or port group view:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the specified VLAN as a multicast VLAN and enter multicast VLAN view.	multicast-vlan <i>vlan-id</i>	By default, a VLAN is not a multicast VLAN.
3. Return to system view.	quit	N/A

Step	Command	Remarks
4. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
5. Configure the current port as a member port of the multicast VLAN.	port multicast-vlan <i>vlan-id</i>	By default, a user port does not belong to any multicast VLAN.

Displaying and maintaining multicast VLAN

Task	Command	Remarks
Display information about a multicast VLAN.	display multicast-vlan [<i>vlan-id</i>] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view

Multicast VLAN configuration examples

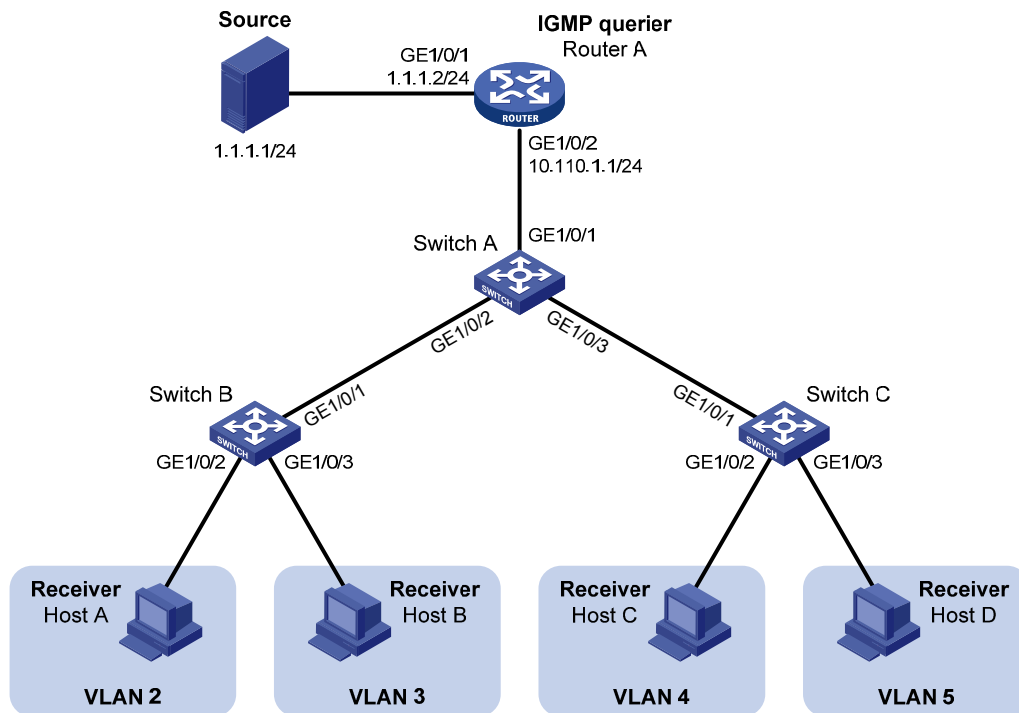
Sub-VLAN-based multicast VLAN configuration example

Network requirements

As shown in [Figure 24](#), IGMPv2 runs on Router A, and IGMPv2 snooping runs on Switch A, Switch B, and Switch C. Router A acts as the IGMP querier. The multicast source sends multicast data to multicast group 224.1.1.1. Host A, Host B, Host C, and Host D are receivers of the multicast group. The hosts belong to VLAN 2 through VLAN 5 respectively.

Configure the sub-VLAN-based multicast VLAN feature on Switch A so that Router A just sends multicast data to Switch A through the multicast VLAN and Switch A forwards the traffic to the receivers that belong to different user VLANs.

Figure 24 Network diagram



Configuration procedure

1. Configure IP addresses:
Configure an IP address and subnet mask for each interface as per Figure 24. (Details not shown.)
2. Configure Router A:
Enable IP multicast routing, enable PIM-DM on each interface and enable IGMP on the host-side interface GigabitEthernet 1/0/2.

```
<RouterA> system-view  
[RouterA] multicast routing-enable  
[RouterA] interface gigabitethernet 1/0/1  
[RouterA-GigabitEthernet1/0/1] pim dm  
[RouterA-GigabitEthernet1/0/1] quit  
[RouterA] interface gigabitethernet 1/0/2  
[RouterA-GigabitEthernet1/0/2] pim dm  
[RouterA-GigabitEthernet1/0/2] igmp enable
```
3. Configure Switch A:
Enable IGMP snooping globally.

```
<SwitchA> system-view  
[SwitchA] igmp-snooping  
[SwitchA-igmp-snooping] quit
```


Create VLAN 2 through VLAN 5.

```
[SwitchA] vlan 2 to 5
```


Configure GigabitEthernet 1/0/2 as a trunk port, and assign it to VLAN 2 and VLAN 3.

```
[SwitchA] interface gigabitethernet 1/0/2  
[SwitchA-GigabitEthernet1/0/2] port link-type trunk
```

```
[SwitchA-GigabitEthernet1/0/2] port trunk permit vlan 2 3
[SwitchA-GigabitEthernet1/0/2] quit
```

Configure GigabitEthernet 1/0/3 as a trunk port, and assign it to VLAN 4 and VLAN 5.

```
[SwitchA] interface gigabitethernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] port link-type trunk
[SwitchA-GigabitEthernet1/0/3] port trunk permit vlan 4 5
[SwitchA-GigabitEthernet1/0/3] quit
```

Create VLAN 10, assign GigabitEthernet 1/0/1 to this VLAN and enable IGMP snooping in the VLAN.

```
[SwitchA] vlan 10
[SwitchA-vlan10] port gigabitethernet 1/0/1
[SwitchA-vlan10] igmp-snooping enable
[SwitchA-vlan10] quit
```

Configure VLAN 10 as a multicast VLAN and configure VLAN 2 through VLAN 5 as its sub-VLANs.

```
[SwitchA] multicast-vlan 10
[SwitchA-mvlan-10] subvlan 2 to 5
[SwitchA-mvlan-10] quit
```

4. Configure Switch B:

Enable IGMP snooping globally.

```
<SwitchB> system-view
[SwitchB] igmp-snooping
[SwitchB-igmp-snooping] quit
```

Create VLAN 2, assign GigabitEthernet 1/0/2 to VLAN 2, and enable IGMP snooping in the VLAN.

```
[SwitchB] vlan 2
[SwitchB-vlan2] port gigabitethernet 1/0/2
[SwitchB-vlan2] igmp-snooping enable
[SwitchB-vlan2] quit
```

Create VLAN 3, assign GigabitEthernet 1/0/3 to VLAN 3, and enable IGMP snooping in the VLAN.

```
[SwitchB] vlan 3
[SwitchB-vlan3] port gigabitethernet 1/0/3
[SwitchB-vlan3] igmp-snooping enable
[SwitchB-vlan3] quit
```

Configure GigabitEthernet 1/0/1 as a trunk port, and assign it to VLAN 2 and VLAN 3.

```
[SwitchB] interface gigabitethernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] port link-type trunk
[SwitchB-GigabitEthernet1/0/1] port trunk permit vlan 2 3
```

5. Configure Switch C:

The configurations on Switch C are similar to those on Switch B.

6. Verify the configuration:

Display information about the multicast VLAN.

```
[SwitchA] display multicast-vlan
Total 1 multicast-vlan(s)
```

```
Multicast vlan 10
  subvlan list:
    vlan 2-5
  port list:
    no port
```

View the IGMP snooping multicast group information on Switch A.

```
[SwitchA] display igmp-snooping group
```

```
Total 5 IP Group(s).
Total 5 IP Source(s).
Total 5 MAC Group(s).
```

Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port

Subvlan flags: R-Real VLAN, C-Copy VLAN

```
Vlan(id):2.
```

```
Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).
Router port(s):total 0 port(s).
IP group(s):the following ip group(s) match to one mac group.
  IP group address:224.1.1.1
    (0.0.0.0, 224.1.1.1):
      Host port(s):total 1 port(s).
        GE1/0/2                (D)
```

```
MAC group(s):
  MAC group address:0100-5e01-0101
    Host port(s):total 1 port(s).
      GE1/0/2
```

```
Vlan(id):3.
```

```
Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).
Router port(s):total 0 port(s).
IP group(s):the following ip group(s) match to one mac group.
  IP group address:224.1.1.1
    (0.0.0.0, 224.1.1.1):
      Host port(s):total 1 port(s).
        GE1/0/2                (D)
```

```
MAC group(s):
  MAC group address:0100-5e01-0101
    Host port(s):total 1 port(s).
      GE1/0/2
```

```
Vlan(id):4.
```

```
Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).
Router port(s):total 0 port(s).
```

```

IP group(s):the following ip group(s) match to one mac group.
  IP group address:224.1.1.1
    (0.0.0.0, 224.1.1.1):
      Host port(s):total 1 port(s).
        GE1/0/3                (D)
MAC group(s):
  MAC group address:0100-5e01-0101
    Host port(s):total 1 port(s).
      GE1/0/3

Vlan(id):5.
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
  Router port(s):total 0 port(s).
IP group(s):the following ip group(s) match to one mac group.
  IP group address:224.1.1.1
    (0.0.0.0, 224.1.1.1):
      Host port(s):total 1 port(s).
        GE1/0/3                (D)
MAC group(s):
  MAC group address:0100-5e01-0101
    Host port(s):total 1 port(s).
      GE1/0/3

Vlan(id):10.
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
  Router port(s):total 1 port(s).
    GE1/0/1                (D)
IP group(s):the following ip group(s) match to one mac group.
  IP group address:224.1.1.1
    (0.0.0.0, 224.1.1.1):
      Host port(s):total 0 port(s).
MAC group(s):
  MAC group address:0100-5e01-0101
    Host port(s):total 0 port(s).

```

The output shows that IGMP snooping is maintaining the router port in the multicast VLAN (VLAN 10) and the member ports in the sub-VLANs (VLAN 2 through VLAN 5).

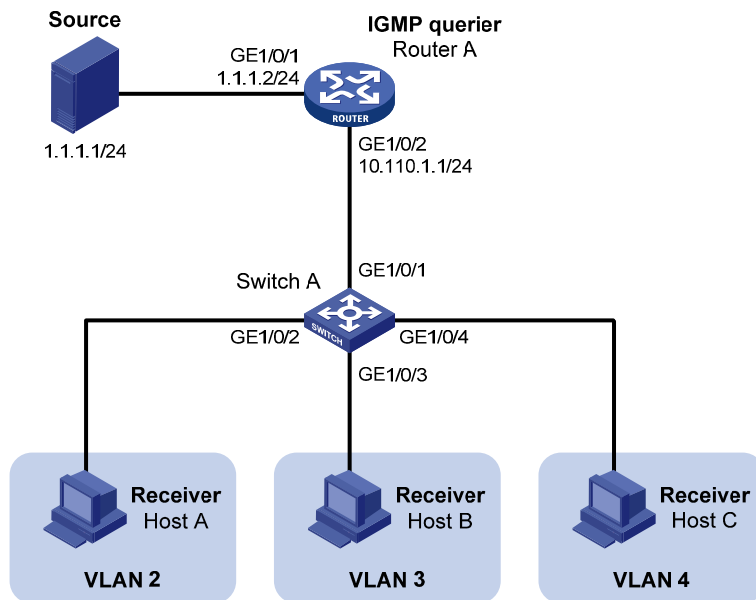
Port-based multicast VLAN configuration example

Network requirements

As shown in [Figure 25](#), IGMPv2 runs on Router A. IGMPv2 Snooping runs on Switch A. Router A acts as the IGMP querier. The multicast source sends multicast data to multicast group 224.1.1.1. Host A, Host B, and Host C are receivers of the multicast group, and the hosts belong to VLAN 2 through VLAN 4 respectively.

Configure the port-based multicast VLAN feature on Switch A so that Router A just sends multicast data to Switch A through the multicast VLAN and Switch A forwards the multicast data to the receivers that belong to different user VLANs.

Figure 25 Network diagram



Configuration procedure

1. Configure IP addresses:
Configure the IP address and subnet mask for each interface as per Figure 25. (Details not shown.)
2. Configure Router A:
Enable IP multicast routing, enable PIM-DM on each interface, and enable IGMP on the host-side interface GigabitEthernet 1/0/2.

```
<RouterA> system-view
[RouterA] multicast routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] pim dm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] pim dm
[RouterA-GigabitEthernet1/0/2] igmp enable
```
3. Configure Switch A:
Enable IGMP snooping globally.

```
<SwitchA> system-view
[SwitchA] igmp-snooping
[SwitchA-igmp-snooping] quit
```

Create VLAN 10, assign GigabitEthernet 1/0/1 to VLAN 10, and enable IGMP snooping in this VLAN.

```
[SwitchA] vlan 10
[SwitchA-vlan10] port gigabitethernet 1/0/1
[SwitchA-vlan10] igmp-snooping enable
[SwitchA-vlan10] quit
```

Create VLAN 2 and enable IGMP snooping in the VLAN.

```
[SwitchA] vlan 2
[SwitchA-vlan2] igmp-snooping enable
[SwitchA-vlan2] quit
```

The configuration for VLAN 3 and VLAN 4 is similar. (Details not shown.)

Configure GigabitEthernet 1/0/2 as a hybrid port. Configure VLAN 2 as the default VLAN. Configure GigabitEthernet 1/0/2 to permit packets of VLAN 2 and VLAN 10 to pass and untag the packets when forwarding them.

```
[SwitchA] interface gigabitethernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] port link-type hybrid
[SwitchA-GigabitEthernet1/0/2] port hybrid pvid vlan 2
[SwitchA-GigabitEthernet1/0/2] port hybrid vlan 2 untagged
[SwitchA-GigabitEthernet1/0/2] port hybrid vlan 10 untagged
[SwitchA-GigabitEthernet1/0/2] quit
```

The configuration for GigabitEthernet 1/0/3 and GigabitEthernet 1/0/4 is similar. (Details not shown.)

Configure VLAN 10 as a multicast VLAN.

```
[SwitchA] multicast-vlan 10
```

Assign GigabitEthernet 1/0/2 and GigabitEthernet 1/0/3 to VLAN 10.

```
[SwitchA-mvlan-10] port gigabitethernet 1/0/2 to gigabitethernet 1/0/3
[SwitchA-mvlan-10] quit
```

Assign GigabitEthernet 1/0/4 to VLAN 10.

```
[SwitchA] interface gigabitethernet 1/0/4
[SwitchA-GigabitEthernet1/0/4] port multicast-vlan 10
[SwitchA-GigabitEthernet1/0/4] quit
```

4. Verify the configuration:

View the multicast VLAN information on Switch A.

```
[SwitchA] display multicast-vlan
Total 1 multicast-vlan(s)
```

```
Multicast vlan 10
  subvlan list:
    no subvlan
  port list:
    GE1/0/2                GE1/0/3                GE1/0/4
```

View the IGMP snooping multicast group information on Switch A.

```
[SwitchA] display igmp-snooping group
Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).
```

Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port

Subvlan flags: R-Real VLAN, C-Copy VLAN

```
Vlan(id):10.
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
```

```
Router port(s):total 1 port(s).
    GE1/0/1                (D)
IP group(s):the following ip group(s) match to one mac group.
IP group address:224.1.1.1
(0.0.0.0, 224.1.1.1):
    Host port(s):total 3 port(s).
        GE1/0/2            (D)
        GE1/0/3            (D)
        GE1/0/4            (D)
MAC group(s):
MAC group address:0100-5e01-0101
    Host port(s):total 3 port(s).
        GE1/0/2
        GE1/0/3
        GE1/0/4
```

The output shows that IGMP snooping is maintaining the router ports and member ports in VLAN 10.

Configuring multicast routing and forwarding (available only on the HP 5500 EI)

Overview

In multicast implementations, the following types of tables implement multicast routing and forwarding:

- **Multicast routing table of a multicast routing protocol**—Each multicast routing protocol has its own multicast routing table, such as PIM routing table.
- **General multicast routing table**—The multicast routing information of different multicast routing protocols forms a general multicast routing table.
- **Multicast forwarding table**—The multicast forwarding table guides the forwarding of multicast packets.

A multicast routing table consists of a set of (S, G) entries. Each entry indicates the routing information for delivering multicast data from a multicast source to a multicast group. If a router supports multiple multicast protocols, its multicast routing table includes routes generated by multiple protocols. The router chooses the optimal route from the multicast routing table based on the configured multicast routing and forwarding policy and adds the route entry to its multicast forwarding table.

The term "router" in this document refers to both routers and Layer 3 switches.

The term "interface" in the multicast routing and forwarding features refers to Layer 3 interfaces, including VLAN interfaces and route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).

RPF check mechanism

A multicast routing protocol relies on the existing unicast routes, MBGP routes, or static multicast routes in creating multicast routing entries. When creating multicast routing table entries, a multicast routing protocol uses the reverse path forwarding (RPF) check mechanism to ensure multicast data delivery along the correct paths. In addition, the RPF check mechanism also helps avoid data loops.

RPF check process

The basis for an RPF check is as follows:

- **Unicast routing table**—Contains the shortest path to each destination subnet.
- **MBGP routing table**—Contains multicast routing information.
- **Static multicast routing table**—Contains the RPF routing information defined by the user through static configuration.

MBGP multicast routing table and static multicast routing table are used for RPF check rather than multicast routing.

When a router performs an RPF check, it searches its unicast routing table, MBGP routing table, and static multicast routing table at the same time. The specific process is as follows:

1. The router chooses an optimal route from the unicast routing table, the MBGP routing table, and the static multicast routing table:
 - The router automatically chooses an optimal unicast route by searching its unicast routing table, and using the IP address of the packet source as the destination address. The outgoing interface in the corresponding routing entry is the RPF interface and the next hop is the RPF neighbor. The router considers the path along which the packet from the RPF neighbor arrived on the RPF interface to be the shortest path that leads back to the source.
 - The router automatically chooses an optimal MBGP route by searching its MBGP routing table, and using the IP address of the packet source as the destination address. The outgoing interface in the corresponding routing entry is the RPF interface and the next hop is the RPF neighbor.
 - The router automatically chooses an optimal static multicast route by searching its static multicast routing table, and using the IP address of the packet source as the destination address. The corresponding routing entry explicitly defines the RPF interface and the RPF neighbor.
2. The router selects one of these optimal routes as the RPF route. The selection process is as follows:
 - If configured to use the longest match principle, the router selects the longest match route from these optimal routes. If the three routes have the same mask, the router selects the route with the highest priority. If the three routes have the same priority, the router selects a route as the RPF route according to the sequence of static multicast route, MBGP route, and unicast route.
 - If not configured to use the longest match principle, the router selects the route with the highest priority. If the three routes have the same priority, the router selects a route as the RPF route according to the sequence of static multicast route, MBGP route, and unicast route.

The "packet source" means different things in different situations:

- For a packet traveling along the shortest path tree (SPT) from the multicast source to the receivers or the rendezvous point (RP), the packet source for RPF check is the multicast source.
- For a packet traveling along the rendezvous point tree (RPT) from the RP to the receivers, or along the source-side RPT from the multicast source to the RP, the packet source for RPF check is the RP.
- For a bootstrap message from the bootstrap router (BSR), the packet source for RPF check is the BSR.

For more information about the concepts of SPT, RPT, source-side RPT, RP, and BSR, see "[Configuring PIM \(available only on the HP 5500 EI\)](#)."

RPF check implementation in multicast

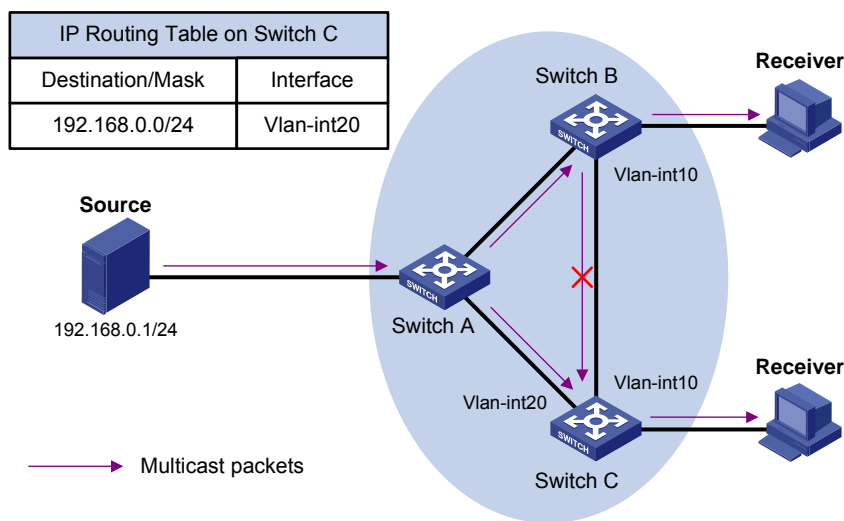
Implementing an RPF check on each received multicast data packet would be a big burden to the router. The use of a multicast forwarding table is the solution to this issue. When creating a multicast routing entry and a multicast forwarding entry for a multicast packet, the router sets the RPF interface of the packet as the incoming interface of the (S, G) entry. After receiving an (S, G) multicast packet, the router first searches its multicast forwarding table:

1. If the corresponding (S, G) entry does not exist in the multicast forwarding table, the packet undergoes an RPF check. The router creates a multicast routing entry based on the relevant routing information and adds the entry into the multicast forwarding table, with the RPF interface as the incoming interface.
 - If the interface that received the packet is the RPF interface, the RPF check succeeds and the router forwards the packet to all the outgoing interfaces.
 - If the interface that received the packet is not the RPF interface, the RPF check fails and the router discards the packet.
2. If the corresponding (S, G) entry exists, and the interface that received the packet is the incoming interface, the router forwards the packet to all the outgoing interfaces.

3. If the corresponding (S, G) entry exists, but the interface that received the packet is not the incoming interface in the multicast forwarding table, the multicast packet undergoes an RPF check.
 - If the RPF interface is the incoming interface of the (S, G) entry, it indicates that the (S, G) entry is correct but the packet arrived from a wrong path. The packet will be discarded.
 - If the RPF interface is not the incoming interface, it indicates that the (S, G) entry has expired, and router replaces the incoming interface with the RPF interface. If the interface on which the packet arrived is the RPF interface, the router forwards the packet to all the outgoing interfaces. Otherwise, it discards the packet.

Assume that unicast routes are available in the network, MBGP is not configured, and no static multicast routes have been configured on Switch C, as shown in [Figure 26](#). Multicast packets travel along the SPT from the multicast source to the receivers. The multicast forwarding table on Switch C contains the (S, G) entry, with VLAN-interface 20 as the incoming interface.

Figure 26 RPF check process



- When a multicast packet arrives on interface VLAN-interface 20 of Switch C, because the interface is the incoming interface of the (S, G) entry, the router forwards the packet to all outgoing interfaces.
- When a multicast packet arrives on interface VLAN-interface 10 of Switch C, because the interface is not the incoming interface of the (S, G) entry, the router performs an RPF check on the packet. The router searches its unicast routing table and finds that the outgoing interface to Source (the RPF interface) is VLAN-interface 20. This means the (S, G) entry is correct, and packet arrived along a wrong path. The RPF check fails and the packet is discarded.

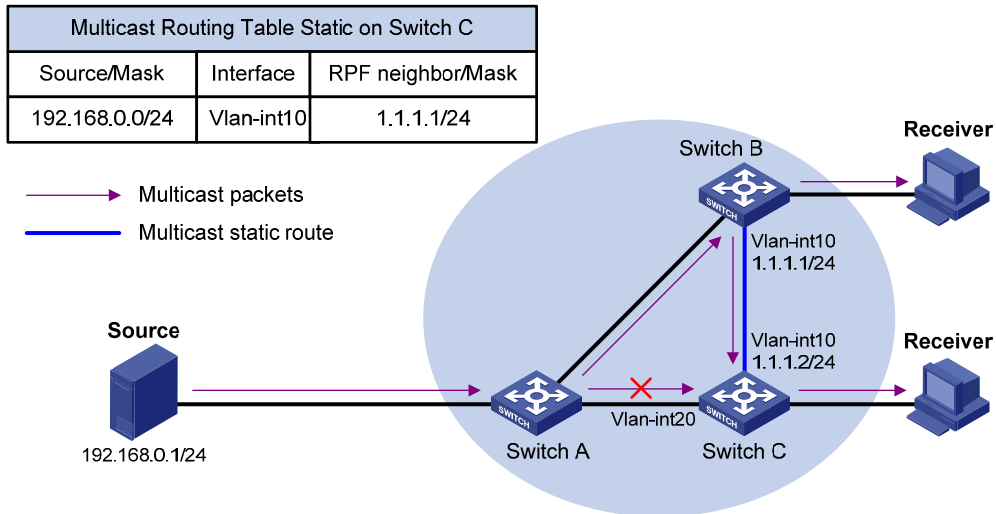
Static multicast routes

A static multicast route is an important basis for RPF check. Depending on the application environment, a static multicast route can change an RPF route and create an RPF route.

Changing an RPF route

Typically, the topology structure of a multicast network is the same as that of a unicast network, and multicast traffic follows the same transmission path as unicast traffic does. You can configure a static multicast route for a given multicast source to change the RPF route to create a transmission path for multicast traffic that is different from that for unicast traffic.

Figure 27 Changing an RPF route

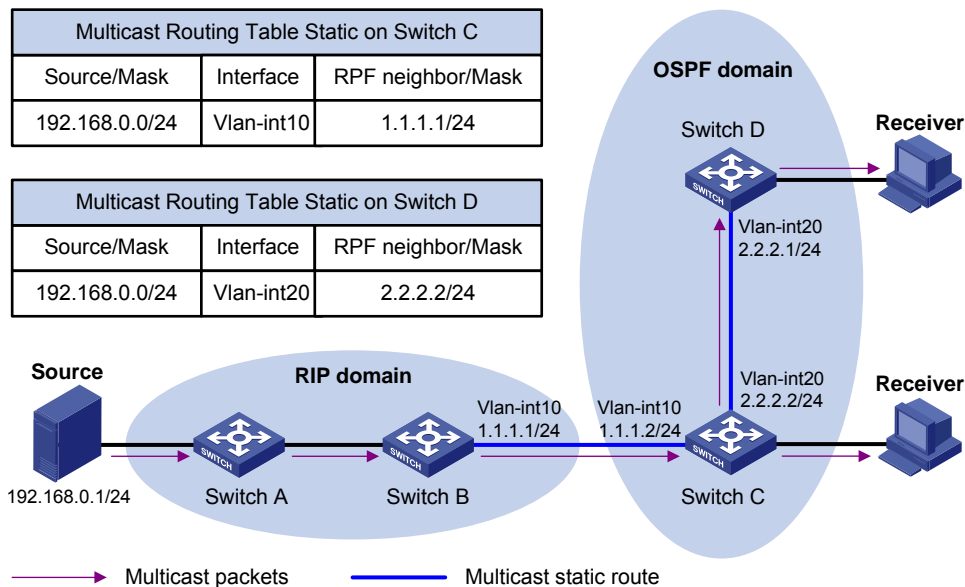


As shown in [Figure 27](#), when no static multicast route is configured, Switch C's RPF neighbor on the path back to Source is Switch A. The multicast information from Source travels along the path from Switch A to Switch C, which is the unicast route between the two routers. When a static multicast route is configured on Switch C and Switch B is configured as Switch C's RPF neighbor on the path back to Source, the multicast information from Source travels from Switch A to Switch B and then to Switch C.

Creating an RPF route

When a unicast route is blocked, multicast traffic forwarding might be stopped because of lack of an RPF route. By configuring a static multicast route for a given multicast source, you can create an RPF route so that a multicast routing entry is created to guide multicast traffic forwarding regardless of whether a unicast route is available.

Figure 28 Creating an RPF route



As shown in [Figure 28](#), the RIP domain and the OSPF domain are unicast isolated from each other. When no static multicast route is configured, the hosts (Receivers) in the OSPF domain cannot receive the multicast packets that the multicast source (Source) sent in the RIP domain. After you configure a static

multicast route on Switch C and Switch D, specifying Switch B as the RPF neighbor of Switch C and specifying Switch C as the RPF neighbor of Switch D, the receivers can receive multicast data that the multicast source sent.

NOTE:

- Static multicast routes only affect RPF check but cannot guide multicast forwarding.
 - A static multicast route is effective only on the multicast router on which it is configured, and will not be advertised throughout the network or redistributed to other routers.
-

Multicast traceroute

You can use the multicast traceroute utility to trace the path that a multicast stream flows down from the first-hop router to the last-hop router.

Concepts in multicast traceroute

- **Last-hop router**—If one of the interfaces of a router connects to the subnet that contains the given destination address, and if the router can forward multicast streams from the given multicast source onto that subnet, that router is called the "last-hop router".
- **First-hop router**—The router that directly connects to the multicast source is called the "first-hop router".
- **Querier**—The router that sends multicast traceroute requests is called the "querier".

Introduction to multicast traceroute packets

A multicast traceroute packet is a special IGMP packet that is different from common IGMP packets in that its IGMP Type field is set to 0x1F or 0x1E and its destination IP address is a unicast address. The following types of multicast traceroute packets are available:

- Query, with the IGMP Type field set to 0x1F,
- Request, with the IGMP Type field set to 0x1F, and
- Response, with the IGMP Type field set to 0x1E.

Process of multicast traceroute

1. The querier sends a query to the last-hop router.
2. After receiving the query, the last-hop router turns the query packet into a request packet by adding a response data block (which contains its interface addresses and packet statistics) to the end of the packet. It then forwards the request packet through unicast to the previous hop for the given multicast source and group.
3. From the last-hop router to the multicast source, each hop adds a response data block to the end of the request packet and unicasts it to the previous hop.
4. When the first-hop router receives the request packet, it changes the packet type to indicate a response packet. Then, it sends the completed packet through unicast to the querier.

Configuration task list

Task	Remarks
Enabling IP multicast routing	Required

Task	Remarks	
Configuring multicast routing and forwarding	Configuring static multicast routes	Optional
	Configuring a multicast routing policy	Optional
	Configuring a multicast forwarding range	Optional
	Configuring the multicast forwarding table size	Optional
	Tracing a multicast path	Optional

! **IMPORTANT:**

IP multicast does not support secondary IP address segments. Namely, multicast can be routed and forwarded only through primary IP addresses even if secondary addresses are configured on the ports. For more information about primary and secondary IP addresses, see *Layer 3— IP Services Configuration Guide*.

Enabling IP multicast routing

Before you configure any Layer 3 multicast functionality, you must enable IP multicast routing.

Enabling IP multicast routing for the public network

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IP multicast routing.	multicast routing-enable	Disabled by default

Enabling IP multicast routing in a VPN instance

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a VPN instance and enter VPN instance view.	ip vpn-instance <i>vpn-instance-name</i>	N/A
3. Configure a route distinguisher (RD) for the VPN instance.	route-distinguisher <i>route-distinguisher</i>	No RD is configured by default.
4. Enable IP multicast routing.	multicast routing-enable	Disabled by default.

For more information about the **ip vpn-instance** and **route-distinguisher** commands, see *IP Routing Command Reference*.

Configuring multicast routing and forwarding

Configuration prerequisites

Before you configure multicast routing and forwarding, complete the following tasks:

- Configure a unicast routing protocol so that all devices in the domain are interoperable at the network layer.
- Enable PIM (PIM-DM or PIM-SM).
- Determine the maximum number of downstream nodes for a single multicast forwarding table entry.
- Determine the maximum number of entries in the multicast forwarding table.

Configuring static multicast routes

By configuring a static multicast route for a given multicast source, you can specify an RPF interface or an RPF neighbor for multicast traffic from that source. If you want to remove a specific static multicast route, use the **undo ip rpf-route-static** command, if you want to remove all static multicast routes, use the **delete ip rpf-route-static** command.

To configure a static multicast route:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure a static multicast route.	ip rpf-route-static [vpn-instance <i>vpn-instance-name</i>] <i>source-address</i> { <i>mask</i> <i>mask-length</i> } [<i>protocol</i> [<i>process-id</i>]] [route-policy <i>policy-name</i>] { <i>rpf-nbr-address</i> <i>interface-type</i> <i>interface-number</i> } [preference <i>preference</i>] [order <i>order-number</i>]	No static multicast route configured by default.
3. Delete static multicast routes.	delete ip rpf-route-static [vpn-instance <i>vpn-instance-name</i>]	Optional.

! IMPORTANT:

When you configure a static multicast route, you cannot specify an RPF neighbor by providing the type and number (*interface-type interface-number*) of the interface if the interface of the RPF neighbor is a Layer 3 Ethernet interface, Layer 3 aggregate interface, Loopback interface, or VLAN interface. Instead, you can specify such an RPF neighbor only by its address (*rpf-nbr-address*).

Configuring a multicast routing policy

You can configure the router to determine the RPF route based on the longest match principle. For more information about RPF route selection, see "[RPF check process](#)."

By configuring per-source or per-source-and-group load splitting, you can optimize the traffic delivery when multiple data flows are handled.

Configuring a multicast routing policy for the public network

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the device to select the RPF route based on the longest match.	multicast longest-match	The route with the highest priority is selected as the RPF route by default.
3. Configure multicast load splitting.	multicast load-splitting { source source-group }	Optional. Disabled by default.

Configuring a multicast routing policy in a VPN instance

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VPN instance view.	ip vpn-instance <i>vpn-instance-name</i>	N/A
3. Configure the device to select the RPF route based on the longest match.	multicast longest-match	The route with the highest priority is selected as the RPF route by default.
4. Configure multicast load splitting.	multicast load-splitting { source source-group }	Optional. Disabled by default.

Configuring a multicast forwarding range

Multicast packets do not travel without a boundary in a network. The multicast data corresponding to each multicast group must be transmitted within a definite scope.

You can configure a forwarding boundary specific to a particular multicast group on all interfaces that support multicast forwarding. A multicast forwarding boundary sets the boundary condition for the multicast groups in the specified range. If the destination address of a multicast packet matches the set boundary condition, the packet will not be forwarded. After you configure an interface as a multicast boundary, the interface can no longer forward multicast packets—including packets sent from the local device—or receive multicast packets.

To configure a multicast forwarding range:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure a multicast forwarding boundary.	multicast boundary <i>group-address</i> { <i>mask</i> <i>mask-length</i> }	No forwarding boundary by default.

Configuring the multicast forwarding table size

Configuration guidelines

The switch maintains the corresponding forwarding entry for each multicast packet that it receives. Excessive multicast routing entries, however, can exhaust the switch's memory and cause lower performance. You can set a limit on the number of entries in the multicast forwarding table based on the networking situation and the performance requirements. If the configured maximum number of multicast forwarding table entries is smaller than the current value, the forwarding entries in excess are not deleted immediately. Instead, the multicast routing protocol that runs on the switch deletes them. The switch no longer adds new multicast forwarding entries until the number of existing multicast forwarding entries comes down below the configured value.

When forwarding multicast traffic, the switch replicates a copy of the multicast traffic for each downstream node and forwards the traffic. Therefore, each of these downstream nodes forms a branch of the multicast distribution tree. You can configure the maximum number of downstream nodes (the

maximum number of outgoing interfaces), for a single entry in the multicast forwarding table to lessen the burden on the switch for replicating multicast traffic. If the configured maximum number of downstream nodes for a single multicast forwarding entry is smaller than the current number, the downstream nodes in excess are not deleted immediately. Instead, the multicast routing protocol that runs on the switch deletes them. The switch no longer adds new multicast forwarding entries for newly added downstream nodes until the number of existing downstream nodes comes down below the configured value.

Configuration procedure

To configure the multicast forwarding table size for the public network:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the maximum number of entries in the multicast forwarding table.	multicast forwarding-table route-limit <i>limit</i>	Optional. 2000 by default.
3. Configure the maximum number of downstream nodes for a single multicast forwarding entry.	multicast forwarding-table downstream-limit <i>limit</i>	Optional. 128 by default.

To configure the multicast forwarding table size in a VPN instance:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VPN instance view.	ip vpn-instance <i>vpn-instance-name</i>	N/A
3. Configure the maximum number of entries in the multicast forwarding table.	multicast forwarding-table route-limit <i>limit</i>	Optional. 2000 by default.
4. Configure the maximum number of downstream nodes for a single route in the multicast forwarding table.	multicast forwarding-table downstream-limit <i>limit</i>	Optional. 128 by default.

Tracing a multicast path

You can run the **mtracert** command to trace the path down which the multicast traffic flows from a given first-hop router to the last-hop router.

To trace a multicast path:

Task	Command	Remarks
Trace a multicast path.	mtracert <i>source-address</i> [[<i>last-hop-router-address</i>] <i>group-address</i>]	Available in any view

Displaying and maintaining multicast routing and forwarding

Task	Command	Remarks
Display multicast boundary information.	display multicast [all-instance vpn-instance <i>vpn-instance-name</i>] boundary [<i>group-address</i> [<i>mask</i> <i>mask-length</i>]] [interface <i>interface-type</i> <i>interface-number</i>] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display multicast forwarding table information.	display multicast [all-instance vpn-instance <i>vpn-instance-name</i>] forwarding-table [<i>source-address</i> [mask { <i>mask</i> <i>mask-length</i> }] <i>group-address</i> [mask { <i>mask</i> <i>mask-length</i> }]] incoming-interface { <i>interface-type</i> <i>interface-number</i> register } outgoing-interface { exclude include match } { <i>interface-type</i> <i>interface-number</i> register } statistics slot <i>slot-number</i>] * [port-info] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display the DF information of the multicast forwarding table .	display multicast [all-instance vpn-instance <i>vpn-instance-name</i>] forwarding-table df-info [<i>rp-address</i>] [slot <i>slot-number</i>] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display information about the multicast routing table.	display multicast [all-instance vpn-instance <i>vpn-instance-name</i>] routing-table [<i>source-address</i> [mask { <i>mask</i> <i>mask-length</i> }] <i>group-address</i> [mask { <i>mask</i> <i>mask-length</i> }]] incoming-interface { <i>interface-type</i> <i>interface-number</i> register } outgoing-interface { exclude include match } { <i>interface-type</i> <i>interface-number</i> register }] * [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display information about the static multicast routing table.	display multicast routing-table [all-instance vpn-instance <i>vpn-instance-name</i>] static [<i>source-address</i> { <i>mask-length</i> <i>mask</i> }] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display RPF route information about the specified multicast source.	display multicast [all-instance vpn-instance <i>vpn-instance-name</i>] rpf-info <i>source-address</i> [<i>group-address</i>] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Clear forwarding entries from the multicast forwarding table.	reset multicast [all-instance vpn-instance <i>vpn-instance-name</i>] forwarding-table { { <i>source-address</i> [mask { <i>mask</i> <i>mask-length</i> }] <i>group-address</i> [mask { <i>mask</i> <i>mask-length</i> }]] incoming-interface { <i>interface-type</i> <i>interface-number</i> register } } * all }	Available in user view
Clear routing entries from the multicast routing table.	reset multicast [all-instance vpn-instance <i>vpn-instance-name</i>] routing-table { { <i>source-address</i> [mask { <i>mask</i> <i>mask-length</i> }] <i>group-address</i> [mask { <i>mask</i> <i>mask-length</i> }]] incoming-interface { <i>interface-type</i> <i>interface-number</i> register } } * all }	Available in user view

! **IMPORTANT:**

The **reset** command clears the information in the multicast routing table or the multicast forwarding table, and thus might cause failure of multicast transmission.

When a routing entry is deleted from the multicast routing table, the corresponding forwarding entry is also deleted from the multicast forwarding table.

When a forwarding entry is deleted from the multicast forwarding table, the corresponding routing entry is also deleted from the multicast routing table.

For more information about designated forwarder (DF), see "[Configuring PIM \(available only on the HP 5500 EI\)](#)."

Configuration examples

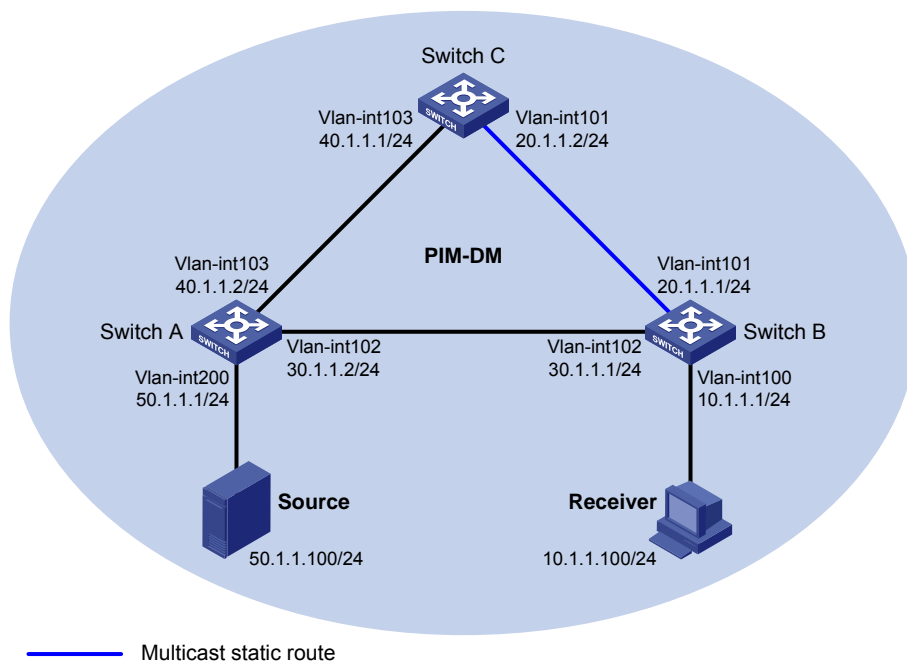
Changing an RPF route

Network requirements

PIM-DM runs in the network. All switches in the network support multicast. Switch A, Switch B, and Switch C run OSPF. Typically, Receiver can receive the multicast data from Source through the path: Switch A to Switch B, which is the same as the unicast route.

Perform the following configuration so that Receiver can receive the multicast data from Source through the path: Switch A to Switch C to Switch B, which is different from the unicast route.

Figure 29 Network diagram



Configuration procedure

1. Configure IP addresses and unicast routing:
Configure the IP address and subnet mask for each interface as per [Figure 29](#). (Details not shown.)

Enable OSPF on the switches in the PIM-DM domain. Ensure the network-layer interoperability among the switches in the PIM-DM domain. Make sure that the switches can dynamically update their routing information by leveraging the unicast routing protocol. (Details not shown.)

2. Enable IP multicast routing, and enable PIM-DM and IGMP:

Enable IP multicast routing on Switch B, enable PIM-DM on each interface, and enable IGMP on VLAN-interface 100.

```
<SwitchB> system-view
[SwitchB] multicast routing-enable
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] igmp enable
[SwitchB-Vlan-interface100] pim dm
[SwitchB-Vlan-interface100] quit
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] pim dm
[SwitchB-Vlan-interface101] quit
[SwitchB] interface vlan-interface 102
[SwitchB-Vlan-interface102] pim dm
[SwitchB-Vlan-interface102] quit
```

Enable IP multicast routing on Switch A, and enable PIM-DM on each interface.

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 200
[SwitchA-Vlan-interface200] pim dm
[SwitchA-Vlan-interface200] quit
[SwitchA] interface vlan-interface 102
[SwitchA-Vlan-interface102] pim dm
[SwitchA-Vlan-interface102] quit
[SwitchA] interface vlan-interface 103
[SwitchA-Vlan-interface103] pim dm
[SwitchA-Vlan-interface103] quit
```

The configuration on Switch C is similar to the configuration on Switch A. (Details not shown.)

Use the **display multicast rpf-info** command to view the RPF route to Source on Switch B.

```
[SwitchB] display multicast rpf-info 50.1.1.100
RPF information about source 50.1.1.100:
  RPF interface: Vlan-interface102, RPF neighbor: 30.1.1.2
  Referenced route/mask: 50.1.1.0/24
  Referenced route type: igp
  Route selection rule: preference-preferred
  Load splitting rule: disable
```

The output shows that the current RPF route on Switch B is contributed by a unicast routing protocol and the RPF neighbor is Switch A.

3. Configure a static multicast route:

Configure a static multicast route on Switch B, specifying Switch C as its RPF neighbor on the route to Source.

```
[SwitchB] ip rpf-route-static 50.1.1.100 24 20.1.1.2
```

4. Verify the configuration:

Use the **display multicast rpf-info** command to view the information about the RPF route to Source on Switch B.

```
[SwitchB] display multicast rpf-info 50.1.1.100
RPF information about source 50.1.1.100:
  RPF interface: Vlan-interface101, RPF neighbor: 20.1.1.2
  Referenced route/mask: 50.1.1.0/24
  Referenced route type: static multicast
  Route selection rule: preference-preferred
  Load splitting rule: disable
```

The output shows that the RPF route on Switch B has changed. It is now the configured static multicast route, and the RPF neighbor is now Switch C.

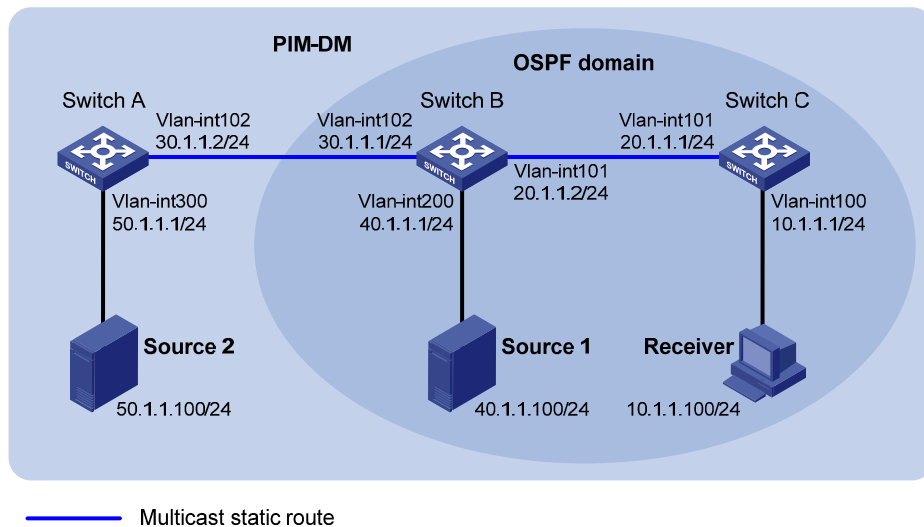
Creating an RPF route

Network requirements

PIM-DM runs in the network and all switches in the network support IP multicast. Switch B and Switch C run OSPF, and have no unicast routes to Switch A. Typically, Receiver can receive the multicast data from Source 1 in the OSPF domain.

Perform the following configuration so that Receiver can receive multicast data from Source 2, which is outside the OSPF domain.

Figure 30 Network diagram



Configuration procedure

1. Configure IP addresses and unicast routing:
Configure the IP address and subnet mask for each interface as per Figure 30. (Details not shown.)
Enable OSPF on Switch B and Switch C. Ensure the network-layer interoperation among Switch B and Switch C. Make sure that the switches can dynamically update their routing information by leveraging the unicast routing protocol. (Details not shown.)
2. Enable IP multicast routing, and enable PIM-DM and IGMP:
Enable IP multicast routing on Switch C, enable PIM-DM on each interface, and enable IGMP on VLAN-interface 100.

```

<SwitchC> system-view
[SwitchC] multicast routing-enable
[SwitchC] interface vlan-interface 100
[SwitchC-Vlan-interface100] igmp enable
[SwitchC-Vlan-interface100] pim dm
[SwitchC-Vlan-interface100] quit
[SwitchC] interface vlan-interface 101
[SwitchC-Vlan-interface101] pim dm
[SwitchC-Vlan-interface101] quit

```

Enable IP multicast routing on Switch A and enable PIM-DM on each interface.

```

<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 300
[SwitchA-Vlan-interface300] pim dm
[SwitchA-Vlan-interface300] quit
[SwitchA] interface vlan-interface 102
[SwitchA-Vlan-interface102] pim dm
[SwitchA-Vlan-interface102] quit

```

The configuration on Switch B is similar to that on Switch A. (Details not shown.)

Use the **display multicast rpf-info** command to view the RPF routes to Source 2 on Switch B and Switch C.

```

[SwitchB] display multicast rpf-info 50.1.1.100
[SwitchC] display multicast rpf-info 50.1.1.100

```

No information is displayed. This means that no RPF route to Source 2 exists on Switch B or Switch C.

3. Configure a static multicast route:

Configure a static multicast route on Switch B, specifying Switch A as its RPF neighbor on the route to Source 2.

```

[SwitchB] ip rpf-route-static 50.1.1.100 24 30.1.1.2

```

Configure a static multicast route on Switch C, specifying Switch B as its RPF neighbor on the route to Source 2.

```

[SwitchC] ip rpf-route-static 10.1.1.100 24 20.1.1.2

```

4. Verify the configuration:

Use the **display multicast rpf-info** command to view the RPF routes to Source 2 on Switch B and Switch C.

```

[SwitchB] display multicast rpf-info 50.1.1.100
RPF information about source 50.1.1.100:
  RPF interface: Vlan-interface102, RPF neighbor: 30.1.1.2
  Referenced route/mask: 50.1.1.0/24
  Referenced route type: static multicast
  Route selection rule: preference-preferred
  Load splitting rule: disable
[SwitchC] display multicast rpf-info 50.1.1.100
RPF information about source 50.1.1.100:
  RPF interface: Vlan-interface101, RPF neighbor: 20.1.1.2
  Referenced route/mask: 50.1.1.0/24
  Referenced route type: static multicast

```

```
Route selection rule: preference-preferred
Load splitting rule: disable
```

The output shows that the RPF routes to Source 2 exist on Switch B and Switch C. The routes are the configured static routes.

Troubleshooting multicast routing and forwarding

Static multicast route failure

Symptom

No dynamic routing protocol is enabled on the routers, and the physical status and link layer status of interfaces are both up, but the static multicast route fails.

Analysis

- If the static multicast route is not configured or updated correctly to match the current network conditions, the route entry and the configuration information of static multicast route do not exist in the multicast routing table.
- If a better route is found, the static multicast route might also fail.

Solution

1. Use the **display multicast routing-table static** command to view the information of static multicast routes to verify that the static multicast route has been correctly configured and that the route entry exists in the multicast routing table.
2. Check the type of the next hop interface of the static multicast route. If the interface is not a point-to-point interface, be sure to specify the next hop address for the outgoing interface when you configure the static multicast route.
3. Check that the static multicast route matches the specified routing protocol. If a protocol was specified in static multicast route configuration, enter the **display ip routing-table** command to check if an identical route was added by the protocol.
4. Check that the static multicast route matches the specified routing policy. If a routing policy was specified when the static multicast route was configured, enter the **display route-policy** command to check the configured routing policy.

Multicast data fails to reach receivers

Symptom

The multicast data can reach some routers but fails to reach the last-hop router.

Analysis

If a multicast forwarding boundary has been configured through the **multicast boundary** command, any multicast packet will be kept from crossing the boundary.

Solution

1. Use the **display pim routing-table** command to verify that the corresponding (S, G) entries exist on the router. If yes, the router has received the multicast data. Otherwise, the router has not received the data.

2. Use the **display multicast boundary** command to check the multicast boundary information on the interfaces. Use the **multicast boundary** command to change the multicast forwarding boundary setting.
3. In the case of PIM-SM, use the **display current-configuration** command to check the BSR and RP information.

Configuring IGMP (available only on the HP 5500 EI)

Overview

As a TCP/IP protocol responsible for IP multicast group member management, the Internet Group Management Protocol (IGMP) is used by IP hosts and adjacent multicast routers to establish and maintain their multicast group memberships.

The term "router" in this document refers to both routers and Layer 3 switches.

The term "interface" in the IGMP features refers to Layer 3 interfaces, including VLAN interfaces and route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).

IGMP versions

- IGMPv1 (defined in RFC 1112)
- IGMPv2 (defined in RFC 2236)
- IGMPv3 (defined in RFC 3376)

All IGMP versions support the Any-Source Multicast (ASM) model. In addition to the ASM model, IGMPv3 can directly implement the Source-Specific Multicast (SSM) model. IGMPv1 and IGMPv2 must work with the IGMP SSM mapping function to implement the SSM model.

For more information about the ASM and SSM models, see "[Multicast overview](#)."

Introduction to IGMPv1

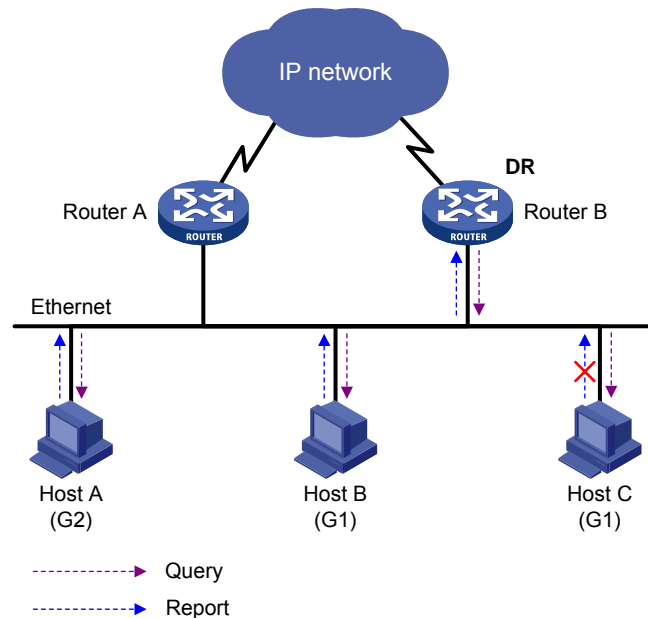
IGMPv1 manages multicast group memberships mainly based on the query and response mechanism.

All multicast routers on the same subnet can receive IGMP membership report messages—often called "reports"—from hosts, but the subnet needs only one router for sending IGMP query messages—often called "queries". The querier election mechanism determines which router acts as the IGMP querier on the subnet.

In IGMPv1, the designated router (DR) elected by the working multicast routing protocol (such as PIM) serves as the IGMP querier.

For more information about DR, see "[Configuring PIM \(available only on the HP 5500 EI\)](#)"

Figure 31 IGMP queries and reports



Assume that Host B and Host C are interested in multicast data addressed to multicast group G1, and Host A is interested in multicast data addressed to G2, as shown in Figure 31. The following process describes how the hosts join the multicast groups and how the IGMP querier (Router B in the figure) maintains the multicast group memberships:

1. The hosts send unsolicited IGMP reports to the addresses of the multicast groups that they want to join, without having to wait for the IGMP queries from the IGMP querier.
2. The IGMP querier periodically multicasts IGMP queries (with the destination address of 224.0.0.1) to all hosts and routers on the local subnet.
3. After receiving a query message, Host B or Host C (the delay timer of whichever expires first) sends an IGMP report to the multicast group address of G1, to announce its membership for G1. Assume that Host B sends the report message. After receiving the report from Host B, Host C (which is on the same subnet as Host B) suppresses its own report for G1, because the IGMP routers (Router A and Router B) have already known that at least one host on the local subnet is interested in G1. This IGMP report suppression mechanism helps reduce traffic on the local subnet.
4. At the same time, because Host A is interested in G2, it sends a report to the multicast group address of G2.
5. Through the query/report process, the IGMP routers determine that members of G1 and G2 are attached to the local subnet, and the multicast routing protocol that is running on the routers (PIM, for example) generates (*, G1) and (*, G2) multicast forwarding entries. These entries will be the basis for subsequent multicast forwarding, where asterisk represents any multicast source.
6. When the multicast data addressed to G1 or G2 reaches an IGMP router, because the (*, G1) and (*, G2) multicast forwarding entries exist on the IGMP router, the router forwards the multicast data to the local subnet, and then the receivers on the subnet receive the data.

IGMPv1 does not specifically define a leave group message (often called a "leave message"). When an IGMPv1 host is leaving a multicast group, it stops sending reports to the address of the multicast group

that it listened to. If no member exists in a multicast group on the subnet, the IGMP router will not receive any report addressed to that multicast group. In this case, the router will delete the multicast forwarding entries for that multicast group after a period of time.

Enhancements in IGMPv2

Compared with IGMPv1, IGMPv2 has introduced a querier election mechanism and a leave-group mechanism.

Querier election mechanism

In IGMPv1, the DR elected by the Layer 3 multicast routing protocol (such as PIM) serves as the querier among multiple routers on the same subnet.

IGMPv2 introduced an independent querier election mechanism. The querier election process is as follows:

1. Initially, every IGMPv2 router assumes itself as the querier and sends IGMP general query messages (often called "general queries") to all hosts and routers on the local subnet. The destination address is 224.0.0.1.
2. After receiving a general query, every IGMPv2 router compares the source IP address of the query message with its own interface address. After comparison, the router with the lowest IP address wins the querier election, and all other IGMPv2 routers become non-queriers.
3. All the non-queriers start a timer, known as "other querier present timer". If a router receives an IGMP query from the querier before the timer expires, it resets this timer. Otherwise, it assumes the querier to have timed out and initiates a new querier election process.

"Leave group" mechanism

In IGMPv1, when a host leaves a multicast group, it does not send any notification to the multicast router. The multicast router relies on the host response timeout timer to determine whether a group has members. This adds to the leave latency.

In IGMPv2, when a host leaves a multicast group, the following steps occur:

1. This host sends a leave message to all routers on the local subnet. The destination address is 224.0.0.2.
2. After receiving the leave message, the querier sends a configurable number of group-specific queries to the group that the host is leaving. The destination address field and group address field of the message are both filled with the address of the multicast group that is being queried.
3. One of the remaining members (if any on the subnet) of the group that is being queried should send a membership report within the maximum response time set in the query messages.
4. If the querier receives a membership report for the group within the maximum response time, it will maintain the memberships of the group. Otherwise, the querier will assume that no hosts on the subnet are still interested in multicast traffic to that group and will stop maintaining the memberships of the group.

Enhancements in IGMPv3

IGMPv3 is based on and is compatible with IGMPv1 and IGMPv2. It provides hosts with enhanced control capabilities and provides enhancements of query and report messages.

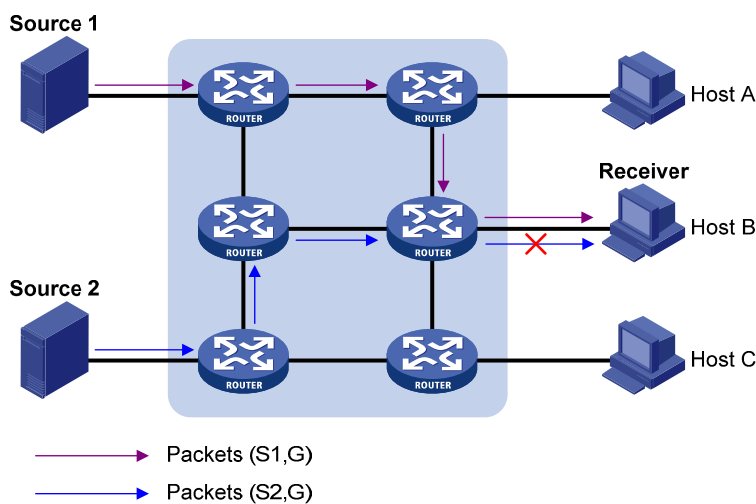
Enhancements in control capability of hosts

IGMPv3 introduced two source filtering modes—Include and Exclude. These modes allow a host to join a designated multicast group and to choose whether to receive or reject multicast data from designated multicast sources. When a host joins a multicast group, one of the following situation occurs:

- If it needs to receive multicast data from specific sources like S1, S2, ..., it sends a report with the Filter-Mode denoted as "Include Sources (S1, S2, ...)".
- If it needs to reject multicast data from specific sources like S1, S2, ..., it sends a report with the Filter-Mode denoted as "Exclude Sources (S1, S2, ...)".

As shown in Figure 32, the network comprises two multicast sources, Source 1 (S1) and Source 2 (S2), both of which can send multicast data to multicast group G. Host B is only interested in the multicast data that Source 1 sends to G but not in the data from Source 2.

Figure 32 Flow paths of source-and-group-specific multicast traffic



In the case of IGMPv1 or IGMPv2, Host B cannot select multicast sources when it joins multicast group G. Therefore, multicast streams from both Source 1 and Source 2 will flow to Host B whether or not it needs them.

When IGMPv3 is running between the hosts and routers, Host B can explicitly express that it needs to receive the multicast data that Source 1 sends to multicast group G—denoted as (S1, G), rather than the multicast data that Source 2 sends to multicast group G—denoted as (S2, G). Thus, only multicast data from Source 1 will be delivered to Host B.

Enhancements in query and report capabilities

1. Query message carrying the source addresses

IGMPv3 supports not only general queries (feature of IGMPv1) and group-specific queries (feature of IGMPv2), but also group-and-source-specific queries.

- A general query does not carry a group address or a source address.
- A group-specific query carries a group address, but no source address.
- A group-and-source-specific query carries a group address and one or more source addresses.

2. Reports containing multiple group records

Unlike an IGMPv1 or IGMPv2 report message, an IGMPv3 report message is destined to 224.0.0.22 and contains one or more group records. Each group record contains a multicast group address and a multicast source address list.

Group records fall into the following categories:

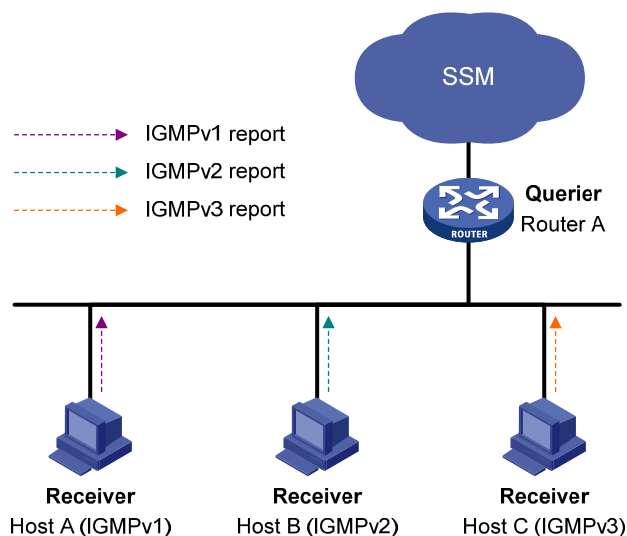
- **IS_IN**—The source filtering mode is Include. The report sender requests the multicast data from only the sources defined in the specified multicast source list.
- **IS_EX**—The source filtering mode is Exclude. The report sender requests the multicast data from any sources but those defined in the specified multicast source list.
- **TO_IN**—The filtering mode has changed from Exclude to Include.
- **TO_EX**—The filtering mode has changed from Include to Exclude.
- **ALLOW**—The Source Address fields in this group record contain a list of the additional sources that the system wants to obtain data from, for packets sent to the specified multicast address. If the change was to an Include source list, these sources are the addresses that were added to the list. If the change was to an Exclude source list, these sources are the addresses that were deleted from the list.
- **BLOCK**—The Source Address fields in this group record contain a list of the sources that the system no longer wants to obtain data from, for packets sent to the specified multicast address. If the change was to an Include source list, these sources are the addresses that were deleted from the list. If the change was to an Exclude source list, these sources are the addresses that were added to the list.

IGMP SSM mapping

The IGMP SSM mapping feature enables you to configure static IGMP SSM mappings on the last-hop router to provide SSM support for receiver hosts that are running IGMPv1 or IGMPv2. The SSM model assumes that the last-hop router has identified the desired multicast sources when receivers join multicast groups.

- When a host that is running IGMPv3 joins a multicast group, it can explicitly specify one or more multicast sources in its IGMPv3 report.
- A host that is running IGMPv1 or IGMPv2, however, cannot specify multicast source addresses in its report. In this case, you must configure the IGMP SSM mapping feature to translate the (*, G) information in the IGMPv1 or IGMPv2 report into (G, INCLUDE, (S1, S2...)) information.

Figure 33 Network diagram



As shown in [Figure 33](#), on an SSM network, Host A, Host B, and Host C are running IGMPv1, IGMPv2, and IGMPv3 respectively. To provide SSM service for all the hosts if IGMPv3 is not available on Host A and Host B, you must configure the IGMP SSM mapping feature on Router A.

With the IGMP SSM mapping feature configured, when Router A receives an IGMPv1 or IGMPv2 report, it checks the multicast group address *G* carried in the message and does the following:

- If *G* is not in the SSM group range, Router A cannot provide the SSM service but can provide the ASM service.
- If *G* is in the SSM group range but no IGMP SSM mappings that correspond to the multicast group *G* have been configured on Router A, Router A cannot provide SSM service and drops the message.
- If *G* is in the SSM group range and the IGMP SSM mappings have been configured on Router A for multicast group *G*, Router A translates the (*, *G*) information in the IGMP report into (*G*, INCLUDE, (S1, S2...)) information based on the configured IGMP SSM mappings and provides SSM service accordingly.

NOTE:

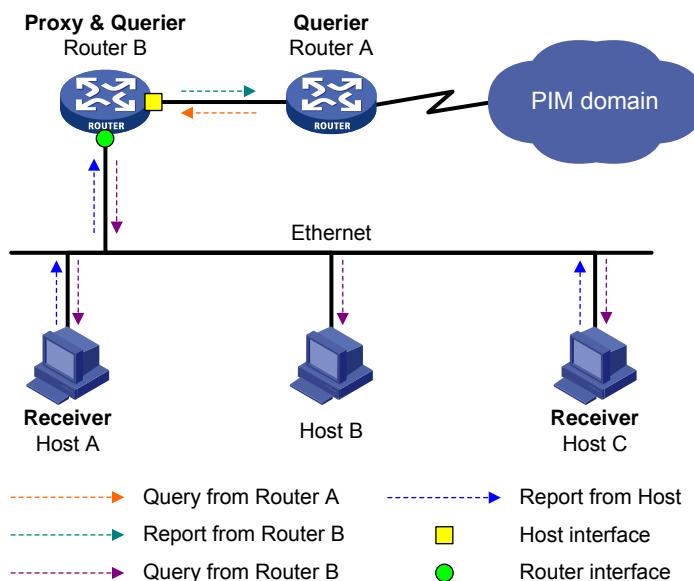
The IGMP SSM mapping feature does not process IGMPv3 reports.

For more information about the SSM group range, see "[Configuring PIM \(available only on the HP 5500 EI\)](#)"

IGMP proxying

In some simple tree-shaped topologies, it is not necessary to configure complex multicast routing protocols, such as PIM, on the boundary devices. Instead, you can configure IGMP proxying on these devices. With IGMP proxying configured, the device serves as a proxy for the downstream hosts to send IGMP messages, maintain group memberships, and implement multicast forwarding based on the memberships. In this case, each boundary device is a host but no longer a PIM neighbor to the upstream device.

Figure 34 Network diagram



As shown in [Figure 34](#), the following types of interfaces are defined on an IGMP proxy device:

- **Upstream interface**—Also called the "proxy interface". A proxy interface is an interface on which IGMP proxying is configured. It is in the direction toward the root of the multicast forwarding tree. An upstream interface acts as a host that is running IGMP. Therefore, it is also called the "host interface".
- **Downstream interface**—An interface that is running IGMP and is not in the direction toward the root of the multicast forwarding tree. A downstream interface acts as a router that is running IGMP. Therefore, it is also called the "router interface".

A device with IGMP proxying configured maintains a group membership database, which stores the group memberships on all the downstream interfaces. Each entry comprises the multicast address, filter mode, and source list. Such an entry is a collection of members in the same multicast group on each downstream interface.

A proxy device performs host functions on the upstream interface based on the database. It responds to queries according to the information in the database or sends join/leave messages when the database changes. On the other hand, the proxy device performs router functions on the downstream interfaces by participating in the querier election, sending queries, and maintaining memberships based on the reports.

IGMP support for VPNs

IGMP maintains group memberships on a per-interface base. After receiving an IGMP message on an interface, IGMP processes the packet within the VPN that the interface belongs to. If IGMP that runs in a VPN needs to exchange information with another multicast protocol, it passes the information only to the protocol that runs in this VPN.

Protocols and standards

- RFC 1112, *Host Extensions for IP Multicasting*
- RFC 2236, *Internet Group Management Protocol, Version 2*
- RFC 3376, *Internet Group Management Protocol, Version 3*
- RFC 4605, *Internet Group Management Protocol (IGMP)/Multicast Listener Discovery (MLD)-Based Multicast Forwarding ("IGMP/MLD Proxying")*

IGMP configuration task list

Task	Remarks	
Configuring basic IGMP functions	Enabling IGMP	Required
	Configuring IGMP versions	Optional
	Configuring static joining	Optional
	Configuring a multicast group filter	Optional
	Setting the maximum number of multicast groups that an interface can join	Optional
Adjusting IGMP performance	Configuring IGMP message options	Optional
	Configuring IGMP query and response parameters	Optional
	Configuring IGMP fast-leave processing	Optional

Task	Remarks	
	Enabling the IGMP host tracking function	Optional
	Setting the DSCP value for IGMP messages	Optional
Configuring IGMP SSM mapping	Enabling SSM mapping	Optional
	Configuring SSM mappings	Optional
Configuring IGMP proxying	Enabling IGMP proxying	Optional
	Configuring multicast forwarding on a downstream interface	Optional

For the configuration tasks in this section:

In IGMP view, the configuration is effective on all interfaces. In interface view, the configuration is effective on only the current interface.

If a feature is not configured on an interface in interface view, the global configuration in IGMP view will apply to that interface. If a feature is configured in both IGMP view and interface view, the configuration in interface view will be given priority.

Configuring basic IGMP functions

Configuration prerequisites

Before you configure basic IGMP functions, complete the following tasks:

- Configure any unicast routing protocol so that all devices in the domain are interoperable at the network layer.
- Configure PIM-DM or PIM-SM.
- Determine the IGMP version.
- Determine the multicast group and multicast source addresses for static group member configuration.
- Determine the ACL rule for multicast group filtering.
- Determine the maximum number of multicast groups that an interface can join.

Enabling IGMP

To configure IGMP, you must enable IGMP on the interface for which the multicast group memberships will be established and maintained.

Enabling IGMP for the public network

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IP multicast routing.	multicast routing-enable	Disabled by default
3. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Enable IGMP.	igmp enable	Disabled by default

Enabling IGMP in a VPN instance

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a VPN instance and enter VPN instance view.	ip vpn-instance <i>vpn-instance-name</i>	N/A
3. Configure an RD for the VPN instance.	route-distinguisher <i>route-distinguisher</i>	No RD is configured by default.
4. Enable IP multicast routing.	multicast routing-enable	Disabled by default.
5. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
6. Bind the interface with a VPN instance.	ip binding vpn-instance <i>vpn-instance-name</i>	By default, an interface belongs to the public network, and is not bound with any VPN instance.
7. Enable IGMP.	igmp enable	Disabled by default.

For more information about the **ip vpn-instance**, **route-distinguisher**, and **ip binding vpn-instance** commands, see *IP Routing Command Reference*.

For more information about the **multicast routing-enable** command, see *IP Multicast Command Reference*.

Configuring IGMP versions

Because the protocol packets of different IGMP versions vary in structure and type, you must configure the same IGMP version for all routers on the same subnet before IGMP can work properly.

Configuring an IGMP version globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network IGMP view or VPN instance IGMP view.	igmp [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure an IGMP version globally.	version <i>version-number</i>	IGMPv2 by default

Configuring an IGMP version on an interface

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure an IGMP version on the interface.	igmp version <i>version-number</i>	IGMPv2 by default

Configuring static joining

Configuration guidelines

After an interface is configured as a static member of a multicast group or a multicast source group, it will act as a virtual member of the multicast group to receive multicast data addressed to that multicast group for the purpose of testing multicast data forwarding.

Before you can configure an interface of a PIM-SM switch as a static member of a multicast group or a multicast source and group, if the interface is PIM-SM enabled, it must be a PIM-SM DR. If the interface is enabled with IGMP but not with PIM-SM, it must be an IGMP querier. For more information about PIM-SM and DR, see "[Configuring PIM \(available only on the HP 5500 EI\)](#)"

A static member port does not respond to queries from the IGMP querier. When you configure a port as a static member port or remove this configuration on the port, the port does not unsolicitedly send any IGMP report or IGMP leave message. In other words, the port is not a real member of the multicast group or the multicast source and group.

Configuration procedure

To configure an interface as a statically connected member of a multicast group or a multicast source and group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the interface as a static member of a multicast group or a multicast source and group.	igmp static-group <i>group-address</i> [source <i>source-address</i>]	An interface is not a static member of any multicast group or multicast source and group by default.

Configuring a multicast group filter

To restrict the hosts on the network attached to an interface from joining certain multicast groups, you can set an ACL rule on the interface as a packet filter so that the interface maintains only the multicast groups the match the criteria.

To configure a multicast group filter:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure a multicast group filter.	igmp group-policy <i>acl-number</i> [<i>version-number</i>]	By default, no multicast group filter is configured on an interface, and hosts on an interface can join any valid multicast group.

Setting the maximum number of multicast groups that an interface can join

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the maximum number of multicast groups that the interface can join.	igmp group-limit <i>limit</i>	2000 by default.

NOTE:

This configuration takes effect for dynamically joined multicast groups but not for statically configured multicast groups.

Adjusting IGMP performance

For the configuration tasks described in this section:

- The configuration in IGMP view is effective on all interfaces, whereas the configuration in interface view is effective only on the current interface.
- If the same feature is configured in both IGMP view and interface view, the configuration in interface view is given priority, regardless of the configuration sequence.

Configuration prerequisites

Before adjusting IGMP performance, complete the following tasks:

- Configure any unicast routing protocol so that all devices in the domain are interoperable at the network layer.
- Configure basic IGMP functions.
- Determine the startup query interval.
- Determine the startup query count.
- Determine the IGMP general query interval.
- Determine the IGMP querier's robustness variable.
- Determine the maximum response time for IGMP general queries.
- Determine the IGMP last-member query interval.
- Determine the other querier present interval.
- Determine the DSCP value for IGMP messages.

Configuring IGMP message options

IGMP queries include group-specific queries and group-and-source-specific queries, and multicast groups change dynamically, so a device cannot maintain the information for all multicast sources and groups. For this reason, when an IGMP router receives a multicast packet but cannot locate the outgoing

interface for the destination multicast group, it must use the Router-Alert option to pass the multicast packet to the upper-layer protocol for processing. For more information about the Router-Alert option, see RFC 2113.

An IGMP message is processed differently depending on whether it carries the Router-Alert option in the IP header:

- By default, for the consideration of compatibility, the switch does not verify the Router-Alert option but processes all the IGMP messages that it received. In this case, IGMP messages are directly passed to the upper-layer protocol, whether or not the IGMP messages carry the Router-Alert option.
- To enhance the switch performance and avoid unnecessary costs, and also for the consideration of protocol security, you can configure the switch to discard IGMP messages that do not carry the Router-Alert option.

Configuring IGMP packet options globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network IGMP view or VPN instance IGMP view.	igmp [<i>vpn-instance</i> <i>vpn-instance-name</i>]	N/A
3. Configure the router to discard any IGMP message that does not carry the Router-Alert option.	require-router-alert	By default, the switch does not check the Router-Alert option.
4. Enable insertion of the Router-Alert option into IGMP messages.	send-router-alert	By default, IGMP messages carry the Router-Alert option.

Configuring IGMP packet options on an interface

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the interface to discard any IGMP message that does not carry the Router-Alert option.	igmp require-router-alert	By default, the switch does not check the Router-Alert option.
4. Enable insertion of the Router-Alert option into IGMP messages.	igmp send-router-alert	By default, IGMP messages carry the Router-Alert option.

Configuring IGMP query and response parameters

IGMP query and response parameters

On startup, the IGMP querier sends IGMP general queries at the startup query interval, which is one-quarter of the IGMP general query interval. The number of queries, or the startup query count, is user configurable.

After startup, the IGMP querier periodically sends IGMP general queries at the IGMP general query interval to check for multicast group members on the network. You can modify the IGMP general query interval based on actual condition of the network.

The IGMPv2 querier sends IGMP group-specific queries at the IGMP last-member query interval when it receives an IGMP leave message. The IGMPv3 querier sends IGMP group-and-source-specific queries at the IGMP last-member query interval when it receives a multicast group and multicast mapping change report. The number of queries, or the last-member query count, equals the robustness variable—the maximum number of packet retransmissions.

A multicast listening host starts a delay timer for each multicast group it has joined when it receives an IGMP query (general query, group-specific query, or group-and-source-specific query). The timer is initialized to a random value in the range of 0 to the maximum response time derived in the IGMP query. When the timer value decreases to 0, the host sends an IGMP report to the corresponding multicast group.

Configuration guidelines

To speed up the response of hosts to IGMP queries and avoid simultaneous timer expirations causing IGMP report traffic bursts, you must properly set the maximum response time.

- For IGMP general queries, the maximum response time is set by the **max-response-time** command.
- For IGMP group-specific queries and IGMP group-and-source-specific queries, the maximum response time equals the IGMP last-member query interval.

When multiple multicast routers exist on the same subnet, the IGMP querier is responsible for sending IGMP queries. If a non-querier router receives no IGMP query from the querier when the other querier present interval expires, it considers that the querier as having failed and starts a new querier election. Otherwise, the non-querier router resets the other querier present timer.

In the configuration, make sure that the other querier present interval is greater than the IGMP general query interval. Otherwise, the IGMP querier might change frequently on the network.

Also make sure that the IGMP general query interval is greater than the maximum response time for IGMP general queries. Otherwise, multicast group members might be wrongly removed.

The configurations of the maximum response time for IGMP general queries, the IGMP last-member query interval and the IGMP other querier present interval are effective only for IGMPv2 and IGMPv3.

Configuration procedure

To configure IGMP query and response parameters globally:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network IGMP view or VPN instance IGMP view.	igmp [vpn-instance vpn-instance-name]	N/A
3. Configure the IGMP querier's robustness variable.	robust-count <i>robust-value</i>	2 by default.
4. Configure the startup query interval.	startup-query-interval <i>interval</i>	By default, the startup query interval is 1/4 of the "IGMP general query interval."
5. Configure the startup query count.	startup-query-count <i>value</i>	By default, the startup query count is set to the IGMP querier's robustness variable.

Step	Command	Remarks
6. Configure the IGMP general query interval.	timer query <i>interval</i>	60 seconds by default.
7. Configure the maximum response time for IGMP general queries.	max-response-time <i>interval</i>	10 seconds by default.
8. Configure the IGMP last-member query interval.	last-member-query-interval <i>interval</i>	1 second by default.
9. Configure the other querier present interval.	timer other-querier-present <i>interval</i>	By default, the other querier present interval is [IGMP general query interval] × [IGMP robustness variable] + [maximum response time for IGMP general queries] / 2.

To configure IGMP query and response parameters on an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the IGMP querier's robustness variable.	igmp robust-count <i>robust-value</i>	2 by default.
4. Configure the startup query interval.	igmp startup-query-interval <i>interval</i>	By default, the startup query interval is 1/4 of the "IGMP general query interval."
5. Configure the startup query count.	igmp startup-query-count <i>value</i>	By default, the startup query count is set to the IGMP querier's robustness variable.
6. Configure the IGMP general query interval.	igmp timer query <i>interval</i>	60 seconds by default.
7. Configure the maximum response time for IGMP general queries.	igmp max-response-time <i>interval</i>	10 seconds by default.
8. Configure the IGMP last-member query interval.	igmp last-member-query-interval <i>interval</i>	1 second by default
9. Configure the other querier present interval.	igmp timer other-querier-present <i>interval</i>	By default, the other querier present interval is [IGMP general query interval] × [IGMP robustness variable] + [maximum response time for IGMP general queries] / 2.

Configuring IGMP fast-leave processing

In some applications, such as ADSL dial-up networking, only one multicast receiver host is attached to a port of the IGMP querier. To allow fast response to the leave messages of the host when it switches

frequently from one multicast group to another, you can enable IGMP fast-leave processing on the IGMP querier.

With fast-leave processing enabled, after receiving an IGMP leave message from a host, the IGMP querier directly sends a leave notification to the upstream without sending IGMP group-specific queries or IGMP group-and-source-specific queries. Thus, the leave latency is reduced on one hand, and the network bandwidth is saved on the other hand.

Configuring IGMP fast-leave processing globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network IGMP view or VPN instance IGMP view.	igmp [vpn-instance vpn-instance-name]	N/A
3. Configure IGMP fast-leave processing.	fast-leave [group-policy acl-number]	Disabled by default

Configuring IGMP fast-leave processing on an interface

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface interface-type interface-number	N/A
3. Configure IGMP fast-leave processing.	igmp fast-leave [group-policy acl-number]	Disabled by default

NOTE:

- The IGMP fast-leave processing configuration is effective only if the switch is running IGMPv2 or IGMPv3.
- The IGMP fast-leave processing configuration is effective on Layer 3 interfaces other than VLAN interfaces, including Layer 3 Ethernet ports, Layer 3 aggregate interfaces, and Tunnel interfaces.

Enabling the IGMP host tracking function

With the IGMP host tracking function, the switch can record the information of the member hosts that are receiving multicast traffic, including the host IP address, running duration, and timeout time. You can monitor and manage the member hosts according to the recorded information.

Enabling the IGMP host tracking function globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network IGMP view/VPN instance IGMP view.	igmp [vpn-instance vpn-instance-name]	N/A

Step	Command	Remarks
3.	Enable the IGMP host tracking function globally. host-tracking	Disabled by default

Enabling the IGMP host tracking function on an interface

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Enter interface view. interface <i>interface-type</i> <i>interface-number</i>	N/A
3.	Enable the IGMP host tracking function on the interface. igmp host-tracking	Disabled by default

Setting the DSCP value for IGMP messages

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Enter public network IGMP view or VPN instance IGMP view. igmp [vpn-instance <i>vpn-instance-name</i>]	N/A
3.	Set the DSCP value for IGMP messages. dscp <i>dscp-value</i>	Optional By default, the DSCP value in IGMP messages is 48.

Configuring IGMP SSM mapping

Because of some possible restrictions, some receiver hosts on an SSM network might run IGMPv1 or IGMPv2. To provide SSM service support for these receiver hosts, configure the IGMP mapping feature on the last-hop router.

Before you configure the IGMP SSM mapping feature, complete the following tasks:

- Configure any unicast routing protocol so that all devices in the domain are interoperable at the network layer.
- Configure basic IGMP functions.

Enabling SSM mapping

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Enter interface view. interface <i>interface-type</i> <i>interface-number</i>	N/A
3.	Enable the IGMP SSM mapping feature. igmp ssm-mapping enable	Disabled by default

NOTE:

To ensure SSM service for all hosts on a subnet, regardless of the IGMP version running on the hosts, enable IGMPv3 on the interface that forwards multicast traffic onto the subnet.

Configuring SSM mappings

By performing this configuration multiple times, you can map a multicast group to different multicast sources.

If IGMPv3 is enabled on a VLAN interface of a switch, and if a port in that VLAN is configured as a simulated host, the simulated host will send IGMPv3 reports even if you did not specify a multicast source when you configure simulated joining with the **igmp-snooping host-join** command. In this case, the corresponding multicast group will not be created based on the configured IGMP SSM mappings. For more information about the **igmp-snooping host-join** command, see *IP Multicast Command Reference*.

To configure an IGMP SSM mapping:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network IGMP view or VPN instance IGMP view.	igmp [vpn-instance vpn-instance-name]	N/A
3. Configure an IGMP SSM mapping.	ssm-mapping group-address { mask mask-length } source-address	No IGMP mappings are configured by default.

Configuring IGMP proxying

Configuration prerequisites

Before you configure the IGMP proxying feature, complete the following tasks:

- Configure any unicast routing protocol so that all devices in the domain are interoperable at the network layer.
- Enable IP multicast routing.

Enabling IGMP proxying

You can enable IGMP proxying on the interface in the direction toward the root of the multicast forwarding tree to make the switch serve as an IGMP proxy.

Configuration guidelines

Each switch can have only one interface serving as the proxy interface. In scenarios with multiple instances, IGMP proxying is configured on only one interface per instance.

You cannot enable IGMP on an interface with IGMP proxying enabled. Moreover, only the **igmp require-router-alert**, **igmp send-router-alert**, and **igmp version** commands can take effect on such an interface.

You cannot enable other multicast routing protocols (such as PIM-DM or PIM-SM) on an interface with IGMP proxying enabled, or vice versa. However, the **source-lifetime**, **source-policy**, and **ssm-policy**

commands configured in PIM view can still take effect. In addition, in IGMPv1, the designated router (DR) is elected by the working multicast routing protocol (such as PIM) to serve as the IGMP querier. Therefore, a downstream interface running IGMPv1 cannot be elected as the DR and thus cannot serve as the IGMP querier.

You cannot enable IGMP proxying on a VLAN interface with IGMP snooping enabled, or vice versa.

Configuration procedure

To enable IGMP proxying:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable the IGMP proxying feature.	igmp proxying enable	Disabled by default

Configuring multicast forwarding on a downstream interface

Only queriers can forward multicast traffic but non-queriers have no multicast forwarding capabilities. This design helps avoid duplicate multicast flows. It is the same on IGMP proxy switches. Only the downstream interfaces acting as a querier can forward multicast traffic to downstream hosts.

However, when a downstream interface of a proxy switch fails to win the querier election, you must enable multicast forwarding on this interface.

ⓘ IMPORTANT:

On a multi-access network with more than one IGMP proxy switch, you cannot enable multicast forwarding on any other non-querier downstream interface after one of the downstream interfaces of these IGMP proxy switches has been elected as the querier. Otherwise, duplicate multicast flows might be received on the multi-access network.

To enable multicast forwarding on a downstream interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable multicast forwarding on a non-querier downstream interface.	igmp proxying forwarding	Disabled by default

Displaying and maintaining IGMP

Task	Command	
Display IGMP group information.	display igmp [all-instance vpn-instance <i>vpn-instance-name</i>] group [<i>group-address</i> interface <i>interface-type interface-number</i>] [static verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the Layer 2 port information of IGMP groups.	display igmp group port-info [vlan <i>vlan-id</i>] [slot <i>slot-number</i>] [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about the hosts tracked by IGMP on an interface.	display igmp host interface <i>interface-type interface-number</i> group <i>group-address</i> [source <i>source-address</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about the hosts tracked by IGMP on the Layer 2 ports.	display igmp host port-info vlan <i>vlan-id</i> group <i>group-address</i> [source <i>source-address</i>] [vlan <i>vlan-id</i>] [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IGMP configuration and operation information.	display igmp [all-instance vpn-instance <i>vpn-instance-name</i>] interface [<i>interface-type interface-number</i>] [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the information of IGMP proxying groups.	display igmp [all-instance vpn-instance <i>vpn-instance-name</i>] proxying group [<i>group-address</i>] [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information in the IGMP routing table.	display igmp [all-instance vpn-instance <i>vpn-instance-name</i>] routing-table [<i>source-address</i> [mask { <i>mask</i> <i>mask-length</i> }] <i>group-address</i> [mask { <i>mask</i> <i>mask-length</i> }] flags { act suc }] * [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IGMP SSM mappings.	display igmp [all-instance vpn-instance <i>vpn-instance-name</i>] ssm-mapping <i>group-address</i> [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	
Display the multicast group information created from IGMPv1 and IGMPv2 reports based on the configured IGMP SSM mappings.	display igmp [all-instance vpn-instance <i>vpn-instance-name</i>] ssm-mapping group [<i>group-address</i> interface <i>interface-type interface-number</i>] [verbose] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display information about the hosts that join the group based on IGMP SSM mappings on an interface.	display igmp ssm-mapping host interface <i>interface-type interface-number</i> group <i>group-address</i> source <i>source-address</i> [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Remove all the dynamic IGMP group entries of IGMP groups.	reset igmp [all-instance vpn-instance <i>vpn-instance-name</i>] group { all interface <i>interface-type interface-number</i> } { all <i>group-address</i> [mask { <i>mask</i> <i>mask-length</i> }] [<i>source-address</i> [mask { <i>mask</i> <i>mask-length</i> }]] }	Available in user view
Remove all the dynamic Layer 2 port entries of IGMP groups.	reset igmp group port-info { all <i>group-address</i> } [vlan <i>vlan-id</i>]	Available in user view
Clear IGMP SSM mappings.	reset igmp [all-instance vpn-instance <i>vpn-instance-name</i>] ssm-mapping group { all interface <i>interface-type interface-number</i> } { all <i>group-address</i> [mask { <i>mask</i> <i>mask-length</i> }] [<i>source-address</i> [mask { <i>mask</i> <i>mask-length</i> }]] }	Available in user view

! **IMPORTANT:**

The **reset igmp group** command might cause an interruption of receivers' reception of multicast data.

The **reset igmp group** command cannot remove static IGMP group entries.

The **display igmp host interface** command can display information about the hosts tracked by IGMP on Layer 3 interfaces other than VLAN interfaces.

The **display igmp ssm-mapping host interface** command can display information about the hosts that join the group based on IGMP SSM mappings on Layer 3 interfaces other than VLAN interfaces.

IGMP configuration examples

Basic IGMP functions configuration example

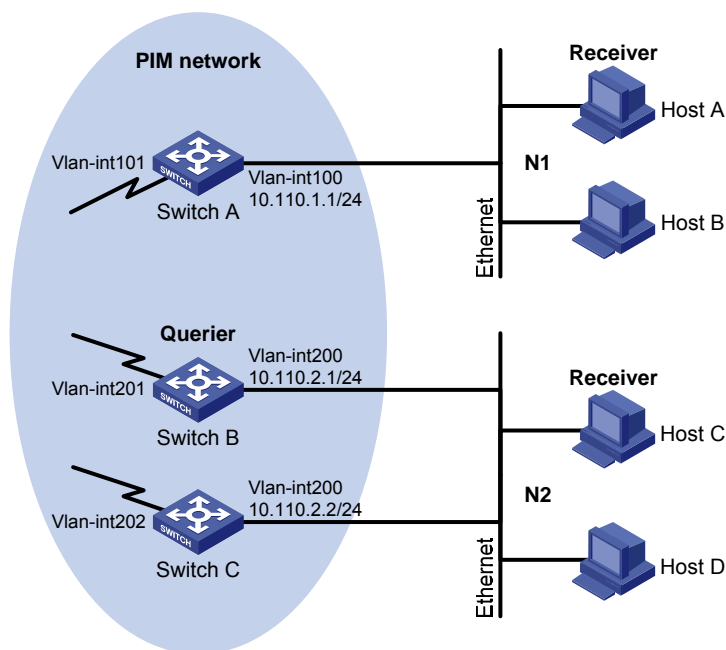
Network requirements

Receivers receive VOD information through multicast. Receivers of different organizations form stub networks N1 and N2, and Host A and Host C are receivers in N1 and N2 respectively.

IGMPv2 runs between Switch A and N1. IGMPv2 runs between the other two switches and N2. Switch B acts as the IGMP querier in N2 because it has a lower IP address.

The hosts in N1 can join only multicast group 224.1.1.1, and the hosts in N2 can join any multicast groups.

Figure 35 Network diagram



Configuration procedure

1. Configure IP addresses and unicast routing:
 - # Configure the IP address and subnet mask of each interface as per Figure 35. (Details not shown.)
 - # Configure the OSPF protocol for interoperation on the PIM network. Ensure the network-layer interoperation on the PIM network and dynamic update of routing information among the switches through a unicast routing protocol. (Details not shown.)
2. Enable IP multicast routing, and enable PIM-DM and IGMP:
 - # Enable IP multicast routing on Switch A, enable PIM-DM on each interface, and enable IGMP on VLAN-interface 100.

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] igmp enable
```

```
[SwitchA-Vlan-interface100] pim dm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim dm
[SwitchA-Vlan-interface101] quit
```

Enable IP multicast routing on Switch B, enable PIM-DM on each interface, and enable IGMP on VLAN-interface 200.

```
<SwitchB> system-view
[SwitchB] multicast routing-enable
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] igmp enable
[SwitchB-Vlan-interface200] pim dm
[SwitchB-Vlan-interface200] quit
[SwitchB] interface vlan-interface 201
[SwitchB-Vlan-interface201] pim dm
[SwitchB-Vlan-interface201] quit
```

Enable IP multicast routing on Switch C, enable PIM-DM on each interface, and enable IGMP on VLAN-interface 200.

```
<SwitchC> system-view
[SwitchC] multicast routing-enable
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] igmp enable
[SwitchC-Vlan-interface200] pim dm
[SwitchC-Vlan-interface200] quit
[SwitchC] interface vlan-interface 202
[SwitchC-Vlan-interface202] pim dm
[SwitchC-Vlan-interface202] quit
```

3. Configure a multicast group filter:

Configure a multicast group filter on Switch A, so that the hosts connected to VLAN-interface 100 can join only multicast group 224.1.1.1.

```
[SwitchA] acl number 2001
[SwitchA-acl-basic-2001] rule permit source 224.1.1.1 0
[SwitchA-acl-basic-2001] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] igmp group-policy 2001
[SwitchA-Vlan-interface100] quit
```

4. Verify the configuration:

Display IGMP information on VLAN-interface 200 of Switch B.

```
[SwitchB] display igmp interface vlan-interface 200
Vlan-interface200(10.110.2.1):
  IGMP is enabled
  Current IGMP version is 2
  Value of query interval for IGMP(in seconds): 60
  Value of other querier present interval for IGMP(in seconds): 125
  Value of maximum query response time for IGMP(in seconds): 10
  Querier for IGMP: 10.110.2.1 (this router)
Total 1 IGMP Group reported
```

SSM mapping configuration example

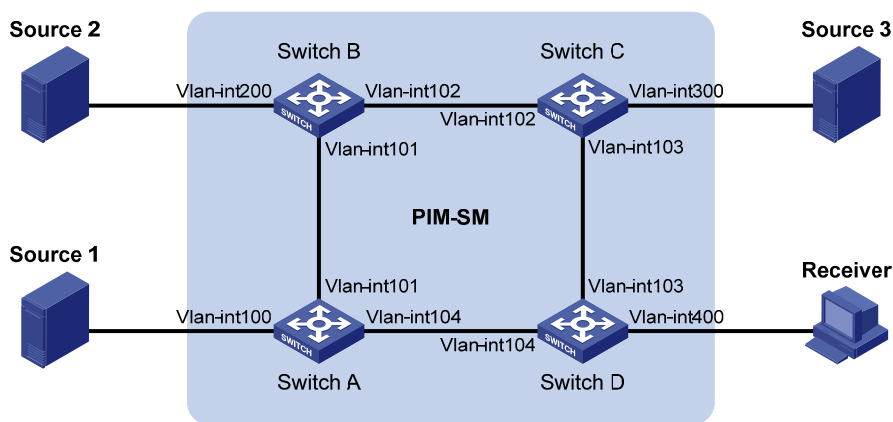
Network requirements

The PIM-SM domain applies both the ASM model and SSM model for multicast delivery. Switch D's VLAN-interface 104 serves as the C-BSR and C-RP. The SSM group range is 232.1.1.0/24.

IGMPv3 runs on Switch D's VLAN-interface 400. The receiver host runs IGMPv2, and does not support IGMPv3. Therefore, the Receiver host cannot specify expected multicast sources in its membership reports.

Source 1, Source 2, and Source 3 send multicast packets to multicast groups in the SSM group range. You can configure the IGMP SSM mapping feature on Switch D so that the receiver host will receive multicast data from Source 1 and Source 3 only.

Figure 36 Network diagram



Device	Interface	IP address	Device	Interface	IP address
Source 1	—	133.133.1.1/24	Source 3	—	133.133.3.1/24
Source 2	—	133.133.2.1/24	Receiver	—	133.133.4.1/24
Switch A	Vlan-int100	133.133.1.2/24	Switch C	Vlan-int300	133.133.3.2/24
	Vlan-int101	192.168.1.1/24		Vlan-int103	192.168.3.1/24
	Vlan-int104	192.168.4.2/24		Vlan-int102	192.168.2.2/24
Switch B	Vlan-int200	133.133.2.2/24	Switch D	Vlan-int400	133.133.4.2/24
	Vlan-int101	192.168.1.2/24		Vlan-int103	192.168.3.2/24
	Vlan-int102	192.168.2.1/24		Vlan-int104	192.168.4.1/24

Configuration procedure

1. Configure IP addresses and unicast routing:
 - # Configure the IP address and subnet mask of each interface as per Figure 36. (Details not shown.)
 - # Configure OSPF for interoperability among the switches. Ensure the network-layer interoperation on the PIM-SM domain and dynamic update of routing information among the switches through a unicast routing protocol. (Details not shown.)
2. Enable IP multicast routing, enable PIM-SM on each interface, and enable IGMP and IGMP SSM mapping on the host-side interface:
 - # Enable IP multicast routing on Switch D, enable PIM-SM on each interface, and enable IGMPv3 and IGMP SSM mapping on VLAN-interface 400.

```

<SwitchD> system-view
[SwitchD] multicast routing-enable
[SwitchD] interface vlan-interface 400
[SwitchD-Vlan-interface400] igmp enable
[SwitchD-Vlan-interface400] igmp version 3
[SwitchD-Vlan-interface400] igmp ssm-mapping enable
[SwitchD-Vlan-interface400] pim sm
[SwitchD-Vlan-interface400] quit
[SwitchD] interface vlan-interface 103
[SwitchD-Vlan-interface103] pim sm
[SwitchD-Vlan-interface103] quit
[SwitchD] interface vlan-interface 104
[SwitchD-Vlan-interface104] pim sm
[SwitchD-Vlan-interface104] quit

```

Enable IP multicast routing on Switch A, and enable PIM-SM on each interface.

```

<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] pim sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim sm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 104
[SwitchA-Vlan-interface104] pim sm
[SwitchA-Vlan-interface104] quit

```

The configuration on Switch B and Switch C is similar to that on Switch A.

3. Configure a C-BSR and a C-RP:

Configure C-BSR and C-RP interfaces on Switch D.

```

[SwitchD] pim
[SwitchD-pim] c-bsr vlan-interface 104
[SwitchD-pim] c-rp vlan-interface 104
[SwitchD-pim] quit

```

4. Configure the SSM group range:

Configure the SSM group range 232.1.1.0/24 on Switch D.

```

[SwitchD] acl number 2000
[SwitchD-acl-basic-2000] rule permit source 232.1.1.0 0.0.0.255
[SwitchD-acl-basic-2000] quit
[SwitchD] pim
[SwitchD-pim] ssm-policy 2000
[SwitchD-pim] quit

```

The configuration on Switch A, Switch B and Switch C is similar to that on Switch D.

5. Configure IGMP SSM mappings:

Configure IGMP SSM mappings on Switch D.

```

[SwitchD] igmp
[SwitchD-igmp] ssm-mapping 232.1.1.0 24 133.133.1.1

```

```
[SwitchD-igmp] ssm-mapping 232.1.1.0 24 133.133.3.1
[SwitchD-igmp] quit
```

6. Verify the configuration:

Display the IGMP SSM mapping information for multicast group 232.1.1.1 on Switch D.

```
[SwitchD] display igmp ssm-mapping 232.1.1.1
Vpn-Instance: public net
Group: 232.1.1.1
Source list:
    133.133.1.1
    133.133.3.1
```

Display the IGMP group information created based on the IGMP SSM mappings on Switch D.

```
[SwitchD] display igmp ssm-mapping group
Total 1 IGMP SSM-mapping Group(s).
Interface group report information of VPN-Instance: public net
Vlan-interface400(133.133.4.2):
  Total 1 IGMP SSM-mapping Group reported
  Group Address      Last Reporter  Uptime      Expires
  232.1.1.1         133.133.4.1   00:02:04    off
```

Display PIM routing table information on Switch D.

```
[SwitchD] display pim routing-table
Vpn-instance: public net
Total 0 (*, G) entry; 2 (S, G) entry
```

```
(133.133.1.1, 232.1.1.1)
  Protocol: pim-ssm, Flag:
  UpTime: 00:13:25
  Upstream interface: Vlan-interface104
    Upstream neighbor: 192.168.4.2
    RPF prime neighbor: 192.168.4.2
  Downstream interface(s) information:
    Total number of downstreams: 1
    1: Vlan-interface400
      Protocol: igmp, UpTime: 00:13:25, Expires: -
```

```
(133.133.3.1, 232.1.1.1)
  Protocol: pim-ssm, Flag:
  UpTime: 00:13:25
  Upstream interface: Vlan-interface103
    Upstream neighbor: 192.168.3.1
    RPF prime neighbor: 192.168.3.1
  Downstream interface(s) information:
    Total number of downstreams: 1
    1: Vlan-interface400
      Protocol: igmp, UpTime: 00:13:25, Expires: -
```

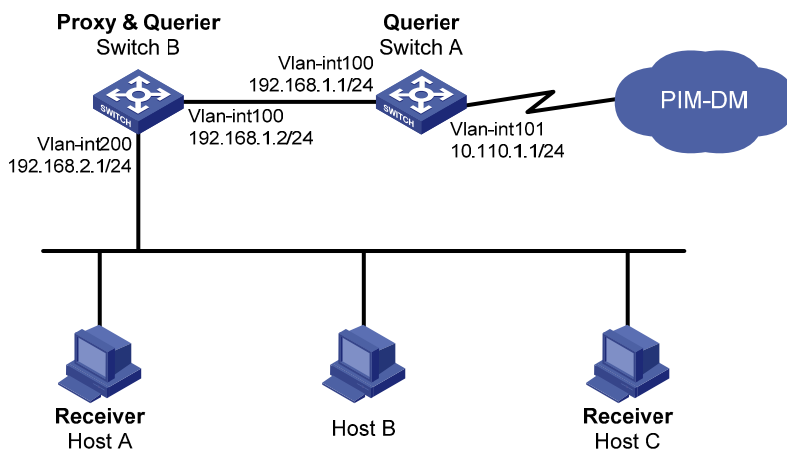

IGMP proxying configuration example

Network requirements

PIM-DM is required to run on the core network. Host A and Host C in the stub network receive VOD information destined to multicast group 224.1.1.1.

Configure the IGMP proxying feature on Switch B so that Switch B can maintain group memberships and forward multicast traffic without running PIM-DM.

Figure 37 Network diagram



Configuration procedure

1. Configure IP addresses:

Configure the IP address and subnet mask of each interface as per Figure 37. (Details not shown.)

2. Enable IP multicast routing, PIM-DM, IGMP, and IGMP proxying:

Enable IP multicast routing on Switch A, PIM-DM on VLAN-interface 101, and IGMP on VLAN-interface 100.

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim dm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] igmp enable
[SwitchA-Vlan-interface100] pim dm
[SwitchA-Vlan-interface100] quit
```

Enable IP multicast routing on Switch B, IGMP Proxying on VLAN-interface 100, and IGMP on VLAN-interface 200.

```
<SwitchB> system-view
[SwitchB] multicast routing-enable
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] igmp proxying enable
[SwitchB-Vlan-interface100] quit
[SwitchB] interface vlan-interface 200
```

```
[SwitchB-Vlan-interface200] igmp enable
[SwitchB-Vlan-interface200] quit
```

3. Verify the configuration:

Display IGMP information on VLAN-interface 100 of Switch B.

```
[SwitchB] display igmp interface vlan-interface 100 verbose
Vlan-interface100(192.168.1.2):
  IGMP proxy is enabled
  Current IGMP version is 2
  Multicast routing on this interface: enabled
  Require-router-alert: disabled
  Version1-querier-present-timer-expiry: 00:00:20
```

Display IGMP group information on Switch A.

```
[SwitchA] display igmp group
Total 1 IGMP Group(s).
Interface group report information of VPN-Instance: public net
Vlan-interface100(192.168.1.1):
  Total 1 IGMP Groups reported
  Group Address      Last Reporter      Uptime             Expires
  224.1.1.1          192.168.1.2       00:02:04           00:01:15
```

The output shows that IGMP reports from the hosts are forwarded to Switch A through the proxy interface, VLAN-interface 100 on Switch B.

Troubleshooting IGMP

No membership information on the receiver-side router

Symptom

When a host sends a report for joining multicast group G, no membership information of the multicast group G exists on the router closest to that host.

Analysis

- The correctness of networking and interface connections and whether the protocol layer of the interface is up directly affect the generation of group membership information.
- Multicast routing must be enabled on the router, and IGMP must be enabled on the interface connecting to the host.
- If the IGMP version on the router interface is lower than that on the host, the router will not be able to recognize the IGMP report from the host.
- If the **igmp group-policy** command has been configured on the interface, the interface cannot receive report messages that fail to pass filtering.

Solution

1. Use the **display igmp interface** command to verify that the networking, interface connection, and IP address configuration are correct. If no information is output, the interface is in an abnormal state. The reason is that you have configured the **shutdown** command on the interface, the interface is not properly connected, or the IP address configuration is not correct.

2. Check that . Use the **display current-configuration** command to verify that multicast routing is enabled. If not, carry out the **multicast routing-enable** command in system view to enable IP multicast routing. In addition, check that IGMP is enabled on the corresponding interfaces.
3. Use the **display igmp interface** command to verify that the IGMP version on the interface is lower than that on the host.
4. Use the **display current-configuration interface** command to verify that no ACL rule has been configured to restrict the host from joining the multicast group G. If the host is restricted from joining the multicast group G, the ACL rule must be modified to allow receiving the reports for the multicast group G.

Inconsistent memberships on routers on the same subnet

Symptom

Different memberships are maintained on different IGMP routers on the same subnet.

Analysis

- A router running IGMP maintains multiple parameters for each interface, and these parameters influence one another, forming very complicated relationships. Inconsistent IGMP interface parameter configurations for routers on the same subnet will surely result in inconsistency of memberships.
- In addition, although an IGMP router is compatible with a host that is running a different version of IGMP, all routers on the same subnet must run the same version of IGMP. Inconsistent IGMP versions running on routers on the same subnet also leads to inconsistency of IGMP memberships.

Solution

1. Use the **display current-configuration** command to verify the IGMP configuration information on the interfaces.
2. Use the **display igmp interface** command on all routers on the same subnet to verify the IGMP-related timer settings. The settings should be consistent on all the routers.
3. Use the **display igmp interface** command to verify that all the routers on the same subnet are running the same version of IGMP.

Configuring PIM (available only on the HP 5500 EI)

PIM overview

Protocol Independent Multicast (PIM) provides IP multicast forwarding by leveraging unicast static routes or unicast routing tables generated by any unicast routing protocol, such as routing information protocol (RIP), open shortest path first (OSPF), intermediate system to intermediate system (IS-IS), or border gateway protocol (BGP). Independent of the unicast routing protocols running on the device, multicast routing can be implemented as long as the corresponding multicast routing entries are created through unicast routes. PIM uses the reverse path forwarding (RPF) mechanism to implement multicast forwarding. When a multicast packet arrives on an interface of the device, it undergoes an RPF check. If the RPF check succeeds, the device creates the corresponding routing entry and forwards the packet. If the RPF check fails, the device discards the packet. For more information about RPF, see "[Configuring multicast routing and forwarding \(available only on the HP 5500 EI\)](#)."

Based on the implementation mechanism, PIM falls into the following categories:

- Protocol Independent Multicast–Dense Mode (PIM-DM)
- Protocol Independent Multicast–Sparse Mode (PIM-SM)
- Bidirectional Protocol Independent Multicast (BIDIR-PIM)
- Protocol Independent Multicast Source-Specific Multicast (PIM-SSM)

The term "router" in this document refers to both routers and Layer 3 switches.

The term "interface" in the PIM features refers to Layer 3 interfaces, including VLAN interfaces and route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).

PIM-DM overview

PIM-DM is a type of dense mode multicast protocol. It uses the push mode for multicast forwarding, and is suitable for small-sized networks with densely distributed multicast members.

The basic implementation of PIM-DM is as follows:

- PIM-DM assumes that at least one multicast group member exists on each subnet of a network. Therefore, multicast data is flooded to all nodes on the network. Then, branches without multicast forwarding are pruned from the forwarding tree, leaving only those branches that contain receivers. This flood-and-prune process takes place periodically. Pruned branches resume multicast forwarding when the pruned state times out. Data is then flooded again down these branches, and then the branches are pruned again.
- When a new receiver on a previously pruned branch joins a multicast group, to reduce the join latency, PIM-DM uses a graft mechanism to resume data forwarding to that branch.

Generally speaking, the multicast forwarding path is a source tree. That is, it is a forwarding tree with the multicast source as its "root" and multicast group members as its "leaves." Because the source tree is the shortest path from the multicast source to the receivers, it is also called a shortest path tree (SPT).

The working mechanism of PIM-DM is summarized as follows:

- Neighbor discovery
- SPT building
- Graft
- Assert

Neighbor discovery

In a PIM domain, a PIM router discovers PIM neighbors, maintains PIM neighboring relationships with other routers, and builds and maintains SPTs by periodically multicasting hello messages to all other PIM routers (224.0.0.13) on the local subnet.

NOTE:

Every PIM-enabled interface on a router sends hello messages periodically, and thus learns the PIM neighboring information pertinent to the interface.

SPT building

The process of building an SPT is the flood-and-prune process.

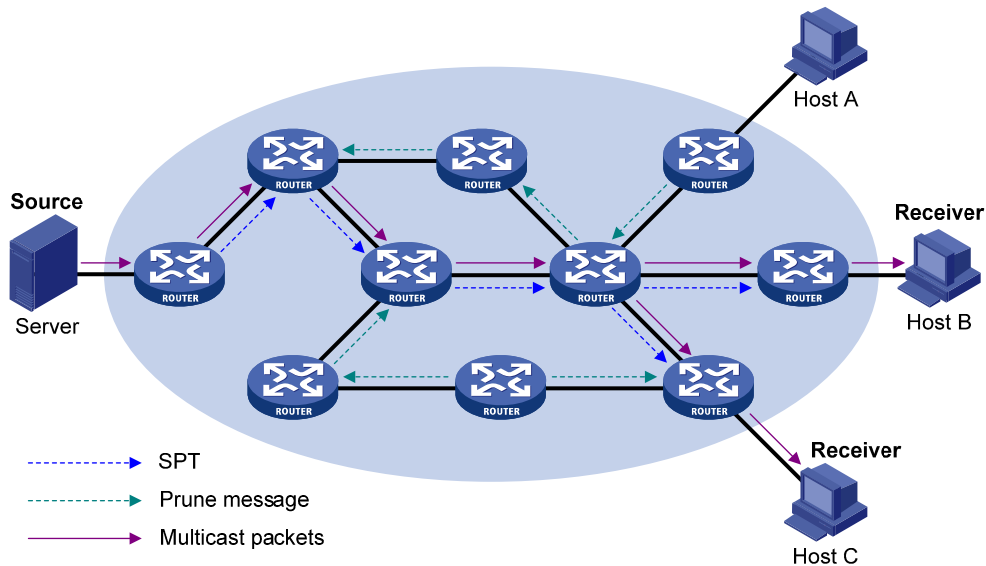
1. In a PIM-DM domain, when a multicast source S sends multicast data to multicast group G , the multicast packet is first flooded throughout the domain. The router first performs RPF check on the multicast packet. If the packet passes the RPF check, the router creates an (S, G) entry and forwards the data to all downstream nodes in the network. In the flooding process, an (S, G) entry is created on all the routers in the PIM-DM domain.
2. Then, nodes without receivers downstream are pruned. A router having no receivers downstream sends a prune message to the upstream node to "tell" the upstream node to delete the corresponding interface from the outgoing interface list in the (S, G) entry and stop forwarding subsequent packets addressed to that multicast group down to this node.

An (S, G) entry contains the multicast source address S , multicast group address G , outgoing interface list, and incoming interface.

For a given multicast stream, the interface that receives the multicast stream is referred to as "upstream," and the interfaces that forward the multicast stream are referred to as "downstream."

A prune process is first initiated by a leaf router. As shown in [Figure 38](#), a router without any receiver attached to it (the router connected with Host A, for example) sends a prune message. This prune process goes on until only necessary branches are left in the PIM-DM domain. These branches constitute the SPT.

Figure 38 SPT building



The flood-and-prune process takes place periodically. A pruned state timeout mechanism is provided. A pruned branch restarts multicast forwarding when the pruned state times out and then is pruned again when it no longer has any multicast receiver.

NOTE:

Pruning has a similar implementation in PIM-SM.

Graft

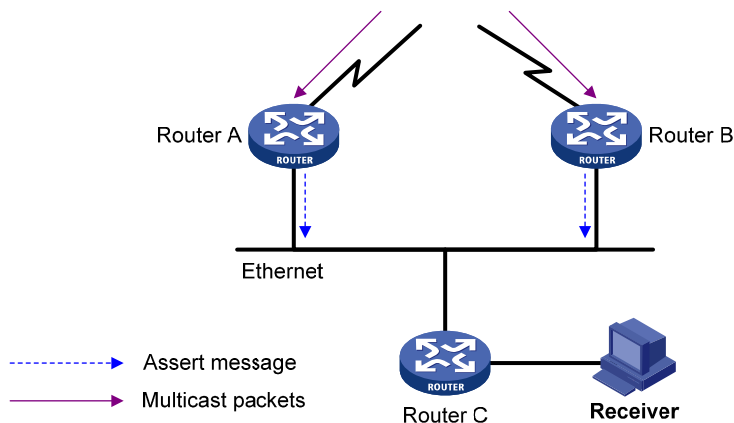
When a host attached to a pruned node joins a multicast group, to reduce the join latency, PIM-DM uses a graft mechanism to resume data forwarding to that branch. The process is as follows:

1. The node that needs to receive multicast data sends a graft message toward its upstream node as a request to join the SPT again.
2. After receiving this graft message, the upstream node puts the interface on which the graft was received into the forwarding state and responds with a graft-ack message to the graft sender.
3. If the node that sent a graft message does not receive a graft-ack message from its upstream node, it will keep sending graft messages at a configurable interval until it receives an acknowledgment from its upstream node.

Assert

Where more than one multicast router exists, the assert mechanism shuts off duplicate multicast flows onto the same multi-access network. It does this by electing a unique multicast forwarder on the multi-access network.

Figure 39 Assert mechanism



As shown in Figure 39, after Router A and Router B receive an (S, G) packet from the upstream node, they both forward the packet to the local subnet. As a result, the downstream node Router C receives two identical multicast packets, and both Router A and Router B, on their own local interface, receive a duplicate packet forwarded by the other. After detecting this condition, both routers send an assert message to all PIM routers (224.0.0.13) on the local subnet through the interface on which the packet was received. The assert message contains the multicast source address (S), the multicast group address (G), and the preference and metric of the unicast route/MBGP route/multicast static route to the source. By comparing these parameters, either Router A or Router B becomes the unique forwarder of the subsequent (S, G) packets on the multi-access subnet. The comparison process is as follows:

1. The router with a higher preference to the source wins;
2. If both routers have the same preference to the source, the router with a smaller metric to the source wins;
3. If a tie exists in route metric to the source, the router with a higher IP address of the local interface wins.

PIM-SM overview

PIM-DM uses the flood-and-prune principle to build SPTs for multicast data distribution. Although an SPT has the shortest path, it is built with a low efficiency. Therefore, the PIM-DM mode is not suitable for large- and medium-sized networks.

PIM-SM is a type of sparse mode multicast protocol. It uses the pull mode for multicast forwarding and is suitable for large-sized and medium-sized networks with sparsely and widely distributed multicast group members.

The basic implementation of PIM-SM is as follows:

- PIM-SM assumes that no hosts need to receive multicast data. In the PIM-SM mode, routers must specifically request a particular multicast stream before the data is forwarded to them. The core task for PIM-SM to implement multicast forwarding will build and maintain rendezvous point trees (RPTs). An RPT is rooted at a router in the PIM domain as the common node, or rendezvous point (RP), through which the multicast data travels along the RPT and reaches the receivers.
- When a receiver is interested in the multicast data addressed to a specific multicast group, the router connected to this receiver sends a join message to the RP that corresponds to that multicast group. The path along which the message goes hop by hop to the RP forms a branch of the RPT.
- When a multicast source sends multicast streams to a multicast group, the source-side designated router (DR) first registers the multicast source with the RP by sending register messages to the RP by

unicast until it receives a register-stop message from the RP. The arrival of a register message at the RP triggers the establishment of an SPT. Then, the multicast source sends subsequent multicast packets along the SPT to the RP. After reaching the RP, the multicast packet is duplicated and delivered to the receivers along the RPT.

NOTE:

Multicast traffic is duplicated only where the distribution tree branches, and this process automatically repeats until the multicast traffic reaches the receivers.

The working mechanism of PIM-SM is summarized as follows:

- Neighbor discovery
- DR election
- RP discovery
- RPT building
- Multicast source registration
- Switchover to SPT
- Assert

Neighbor discovery

PIM-SM uses a similar neighbor discovery mechanism as PIM-DM does. For more information, see "[Neighbor discovery](#)."

DR election

PIM-SM also uses hello messages to elect a DR for a multi-access network (such as Ethernet). The elected DR will be the only multicast forwarder on this multi-access network.

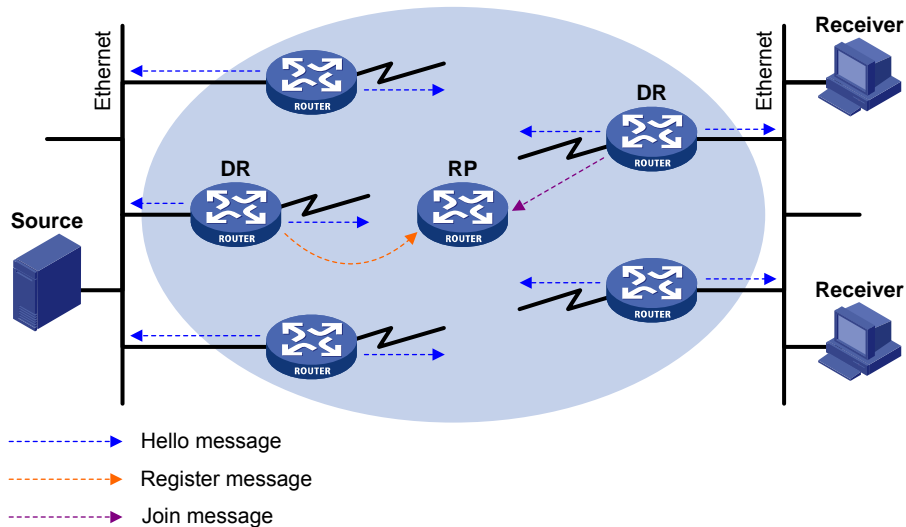
A DR must be elected in a multi-access network, no matter this network connects to multicast sources or to receivers. The receiver-side DR sends join messages to the RP. The source-side DR sends register messages to the RP.

A DR is elected on a multi-access subnet by means of comparison of the priorities and IP addresses carried in hello messages. An elected DR is substantially meaningful to PIM-SM. PIM-DM itself does not require a DR. However, if IGMPv1 runs on any multi-access network in a PIM-DM domain, a DR must be elected to act as the IGMPv1 querier on that multi-access network.

IGMP must be enabled on a device that acts as a receiver-side DR before receivers attached to this device can join multicast groups through this DR.

For more information about IGMP, see "[Configuring IGMP \(available only on the HP 5500 EI\)](#)."

Figure 40 DR election



As shown in Figure 40, the DR election process is as follows:

1. Routers on the multi-access network send hello messages to one another. The hello messages contain the router priority for DR election. The router with the highest DR priority will become the DR.
2. In the case of a tie in the router priority, or if any router in the network does not support carrying the DR-election priority in hello messages, the router with the highest IP address will win the DR election.

When the DR fails, a timeout in receiving a hello message triggers a new DR election process among the other routers.

RP discovery

The RP is the core of a PIM-SM domain. For a small-sized, simple network, one RP is enough for forwarding information throughout the network, and you can statically specify the position of the RP on each router in the PIM-SM domain. In most cases, however, a PIM-SM network covers a wide area and a huge amount of multicast traffic must be forwarded through the RP. To lessen the RP burden and optimize the topological structure of the RPT, you can configure multiple candidate-RPs (C-RPs) in a PIM-SM domain, among which an RP is dynamically elected through the bootstrap mechanism. Each elected RP serves a different multicast group range. For this purpose, you must configure a bootstrap router (BSR). The BSR serves as the administrative core of the PIM-SM domain. A PIM-SM domain can have only one BSR, but can have multiple candidate-BSRs (C-BSRs). If the BSR fails, a new BSR is automatically elected from the C-BSRs to avoid service interruption.

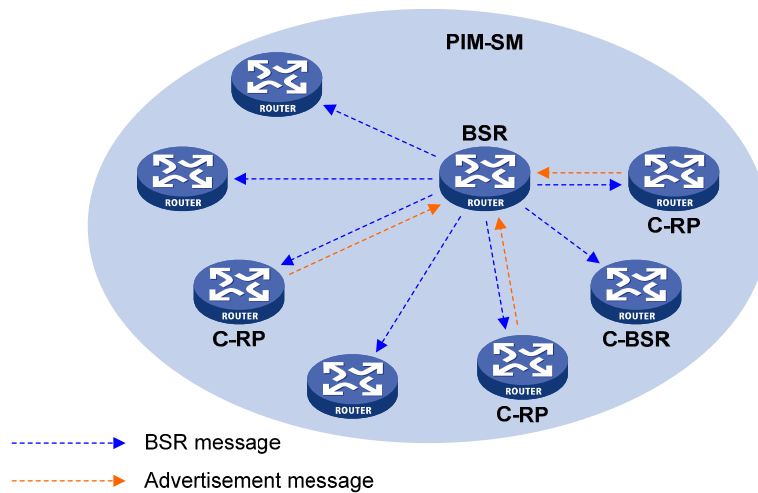
NOTE:

- An RP can serve multiple multicast groups or all multicast groups. Only one RP can serve a given multicast group at a time.
 - A device can serve as a C-RP and a C-BSR at the same time.
-

As shown in Figure 41, each C-RP periodically unicasts its advertisement messages (C-RP-Adv messages) to the BSR. A C-RP-Adv message contains the address of the advertising C-RP and the multicast group range that it serves. The BSR collects these advertisement messages and chooses the appropriate C-RP information for each multicast group to form an RP-set, which is a database of mappings between

multicast groups and RPs. The BSR then encapsulates the RP-set in the bootstrap messages (BSMs) that it periodically originates and floods the bootstrap messages to the entire PIM-SM domain.

Figure 41 BSR and C-RPs



Based on the information in the RP-sets, all routers in the network can calculate the location of the corresponding RPs based on the following rules:

1. The C-RP with the highest priority wins.
2. If all the C-RPs have the same priority, their hash values are calculated through the hashing algorithm. The C-RP with the largest hash value wins.
3. If all the C-RPs have the same priority and hash value, the C-RP that has the highest IP address wins.

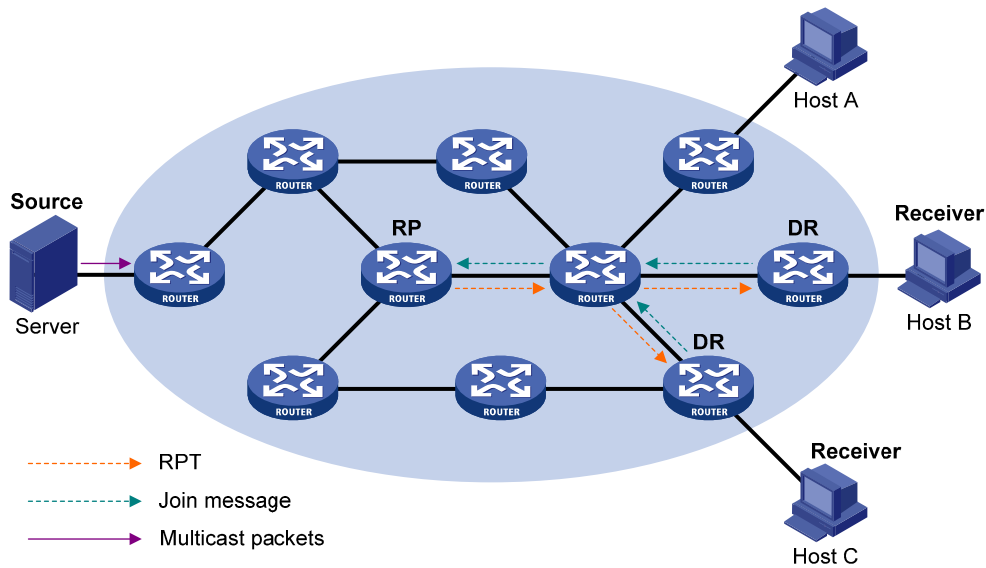
The hashing algorithm used for RP calculation is "Value (G, M, C_i) = (1103515245 * ((1103515245 * (G & M) + 12345) XOR C_i) + 12345) mod 2³¹."

Values in the hashing algorithm

Value	Description
Value	Hash value
G	IP address of the multicast group
M	Hash mask length
C _i	IP address of the C-RP
&	Logical operator of "and"
XOR	Logical operator of "exclusive-or"
Mod	Modulo operator, which gives the remainder of an integer division

RPT building

Figure 42 RPT building in a PIM-SM domain



As shown in [Figure 42](#), the process of building an RPT is as follows:

1. When a receiver joins multicast group G , it uses an IGMP message to inform the directly connected DR.
2. After getting the receiver information, the DR sends a join message, which is forwarded hop by hop to the RP that corresponds to the multicast group.
3. The routers along the path from the DR to the RP form an RPT branch. Each router on this branch generates a $(*, G)$ entry in its forwarding table. The asterisk means any multicast source. The RP is the root of the RPT, and the DRs are the leaves of the RPT.

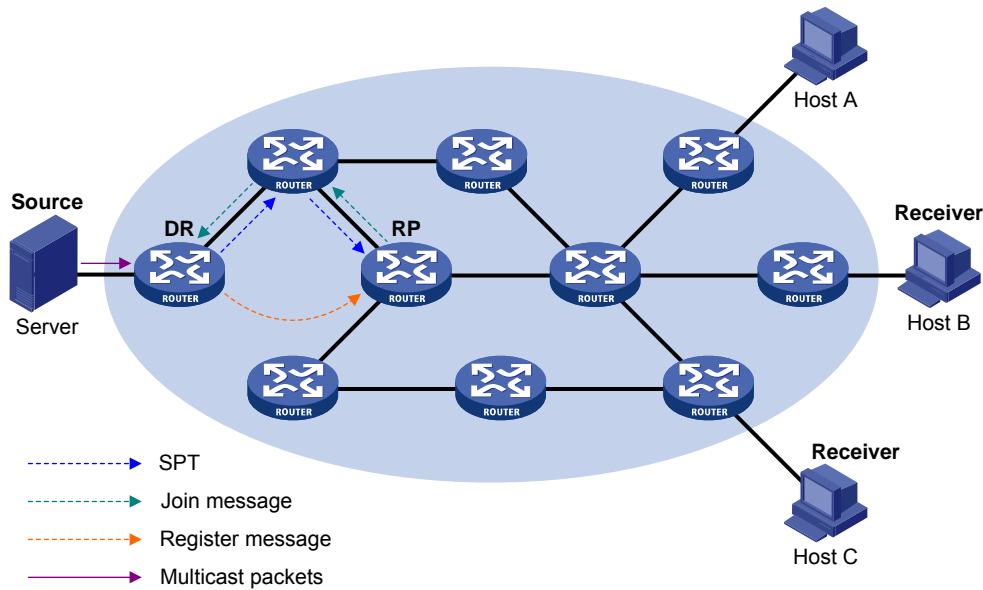
The multicast data addressed to the multicast group G flows through the RP, reaches the corresponding DR along the established RPT, and finally is delivered to the receiver.

When a receiver is no longer interested in the multicast data addressed to multicast group G , the directly connected DR sends a prune message, which goes hop by hop along the RPT to the RP. After receiving the prune message, the upstream node deletes the interface that connects to this downstream node from the outgoing interface list and determines whether it has receivers for that multicast group. If not, the router continues to forward the prune message to its upstream router.

Multicast source registration

The purpose of multicast source registration will inform the RP about the existence of the multicast source.

Figure 43 Multicast source registration



As shown in Figure 43, the multicast source registers with the RP as follows:

1. The multicast source S sends the first multicast packet to multicast group G. After receiving the multicast packet, the DR that directly connects to the multicast source encapsulates the packet in a PIM register message. Then it sends the message to the corresponding RP by unicast.
2. When the RP receives the register message, it extracts the multicast packet from the register message and forwards the multicast packet down the RPT, and sends an (S, G) join message hop by hop toward the multicast source. Thus, the routers along the path from the RP to the multicast source constitute an SPT branch. Each router on this branch generates an (S, G) entry in its forwarding table. The source-side DR is the root of the SPT, and the RP is the leaf of the SPT.
3. The subsequent multicast data from the multicast source travels along the established SPT to the RP. Then the RP forwards the data along the RPT to the receivers. When the multicast traffic arrives at the RP along the SPT, the RP sends a register-stop message to the source-side DR by unicast to stop the source registration process.

NOTE:

The RP is configured to initiate an SPT switchover as described in this section. Otherwise, the source-side DR keeps encapsulating multicast data in register messages, and the registration process will not stop unless no outgoing interfaces exist in the (S, G) entry on the RP.

Switchover to SPT

In a PIM-SM domain, a multicast group corresponds to one RP and RPT. Before the SPT switchover occurs, the source-side DR encapsulates all multicast data destined to the multicast group in register messages and sends these messages to the RP. After receiving these register messages, the RP extracts the multicast data and sends the multicast data down the RPT to the DRs at the receiver side. The RP acts as a transfer station for all multicast packets. The whole process involves the following issues:

- The source-side DR and the RP need to implement complicated encapsulation and de-encapsulation of multicast packets.
- Multicast packets are delivered along a path that might not be the shortest one.
- An increase in multicast traffic adds a great burden on the RP, increasing the risk of failure.

To solve the issues, PIM-SM allows an RP or the DR at the receiver side to initiate an SPT switchover process.

1. The RP initiates an SPT switchover process.

The RP can periodically check the passing-by IPv4 multicast packets. If it finds that the traffic rate exceeds a configurable threshold, the RP sends an (S, G) join message hop by hop toward the multicast source to establish an SPT between the DR at the source side and the RP. Subsequent multicast data travels along the established SPT to the RP.

For more information about the SPT switchover initiated by the RP, see "[Multicast source registration](#)."

2. The receiver-side DR initiates an SPT switchover process.

After receiving the first multicast packet, the receiver-side DR initiates an SPT switchover process, as follows:

- The receiver-side DR sends an (S, G) join message hop by hop toward the multicast source. When the join message reaches the source-side DR, all the routers on the path have installed the (S, G) entry in their forwarding table, and thus an SPT branch is established.
- When the multicast packets travel to the router where the RPT and the SPT deviate, the router drops the multicast packets received from the RPT and sends an RP-bit prune message hop by hop to the RP. After receiving this prune message, the RP sends a prune message toward the multicast source (suppose only one receiver exists). Thus, SPT switchover is completed.
- Multicast data is directly sent from the source to the receivers along the SPT.

PIM-SM builds SPTs through SPT switchover more economically than PIM-DM does through the flood-and-prune mechanism.

Assert

PIM-SM uses a similar assert mechanism as PIM-DM does. For more information, see "[Assert](#)."

BIDIR-PIM overview

In some many-to-many applications, such as multi-side video conference, there might be multiple receivers interested in multiple multicast sources simultaneously. With PIM-DM or PIM-SM, each router along the SPT must create an (S, G) entry for each multicast source, consuming a lot of system resources.

BIDIR-PIM addresses the problem. Derived from PIM-SM, BIDIR-PIM builds and maintains bidirectional RPTs, each of which is rooted at an RP and connects multiple multicast sources with multiple receivers. Traffic from the multicast sources is forwarded through the RPs to the receivers along the bidirectional RPTs. Each router needs to maintain only one (*, G) multicast routing entry, saving system resources.

BIDIR-PIM is suitable for networks with dense multicast sources and dense receivers.

The working mechanism of BIDIR-PIM is summarized as follows:

- Neighbor discovery
- RP discovery
- DF election
- Bidirectional RPT building

Neighbor discovery

BIDIR-PIM uses the same neighbor discovery mechanism as PIM-SM does. For more information, see "[Neighbor discovery](#)."

RP discovery

BIDIR-PIM uses the same RP discovery mechanism as PIM-SM does. For more information, see "[RP discovery](#)."

In PIM-SM, an RP must be specified with a real IP address. In BIDIR-PIM, however, an RP can be specified with a virtual IP address, which is called the rendezvous point address (RPA). The link corresponding to the RPA's subnet is called the rendezvous point link (RPL). All interfaces connected to the RPL can act as the RP, and they back up one another.

NOTE:

In BIDIR-PIM, an RPF interface is the interface pointing to an RP, and an RPF neighbor is the address of the next hop to the RP.

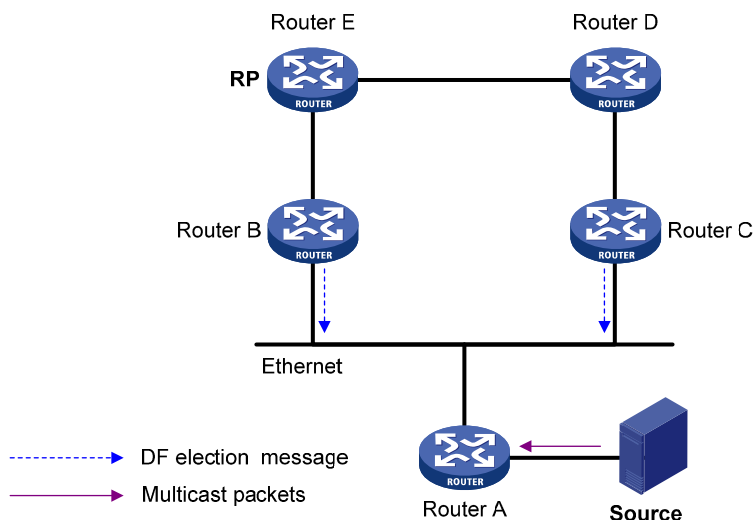
DF election

On a network segment with multiple multicast routers, the same multicast packets might be forwarded to the RP repeatedly. To address this issue, BIDIR-PIM uses a DF election mechanism to elect a unique designated forwarder (DF) for each RP on every network segment within the BIDIR-PIM domain, and allows only the DF to forward multicast data to the RP.

NOTE:

DF election is not necessary for an RPL.

Figure 44 DF election



As shown in [Figure 44](#), without the DF election mechanism, both Router B and Router C can receive multicast packets from Router A, and they might both forward the packets to downstream routers on the local subnet. As a result, the RP (Router E) receives duplicate multicast packets. With the DF election mechanism, once receiving the RP information, Router B and Router C initiate a DF election process for the RP:

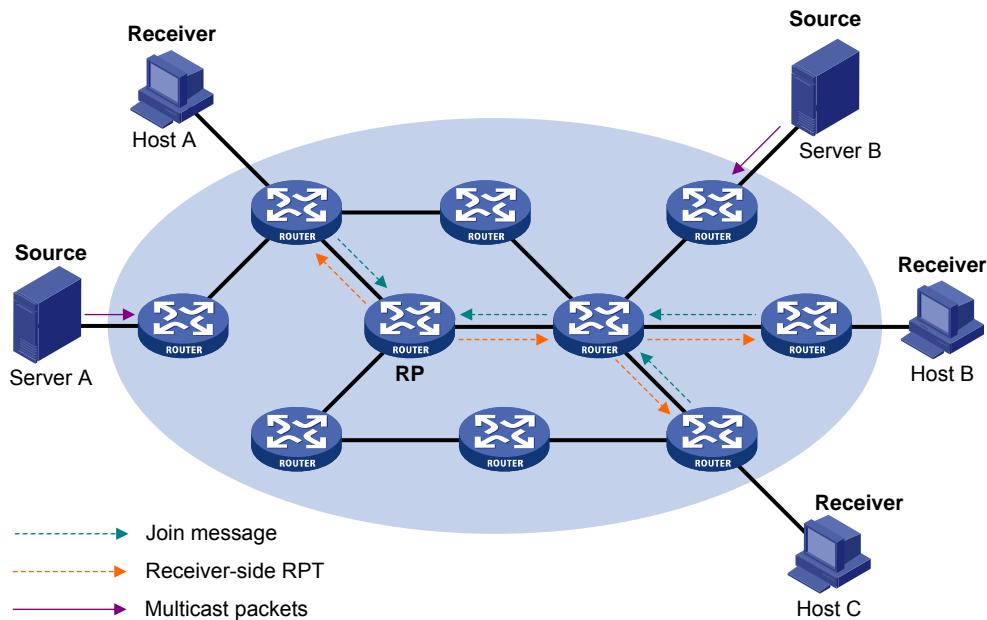
1. Router B and Router C multicast DF election messages to all PIM routers (224.0.0.13). The election messages carry the RP's address, and the priority and metric of the unicast route, MBGP route, or multicast static route to the RP.
2. The router with a route of the highest priority becomes the DF.
3. In the case of a tie, the router with the route with the lowest metric wins the DF election.

4. In the case of a tie in the metric, the router with the highest IP address wins.

Bidirectional RPT building

A bidirectional RPT comprises a receiver-side RPT and a source-side RPT. The receiver-side RPT is rooted at the RP and takes the routers directly connected to the receivers as leaves. The source-side RPT is also rooted at the RP but takes the routers directly connected to the sources as leaves. The processes for building these two parts are different.

Figure 45 RPT building at the receiver side

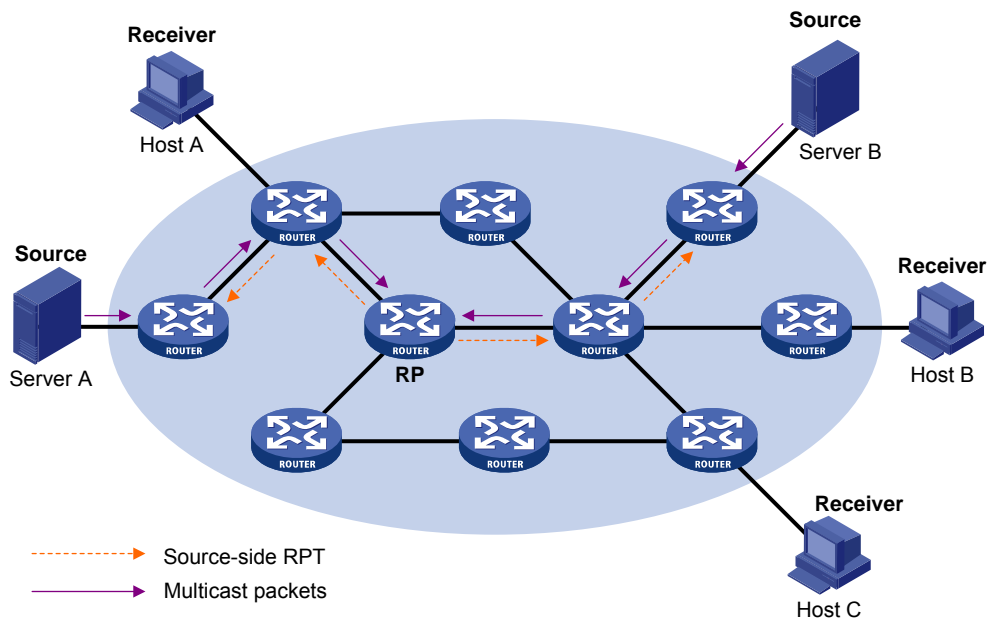


As shown in [Figure 45](#), the process for building a receiver-side RPT is similar to that for building an RPT in PIM-SM:

1. When a receiver joins multicast group G , it uses an IGMP message to inform the directly connected router.
2. After getting the receiver information, the router sends a join message, which is forwarded hop by hop to the RP of the multicast group.
3. The routers along the path from the receiver's directly connected router to the RP form an RPT branch, and each router on this branch adds a $(*, G)$ entry to its forwarding table. The $*$ means any multicast source.

When a receiver is no longer interested in the multicast data addressed to multicast group G , the directly connected router sends a prune message, which goes hop by hop along the reverse direction of the RPT to the RP. After receiving the prune message, each upstream node deletes the interface connected to the downstream node from the outgoing interface list and checks whether it has receivers in that multicast group. If not, the router continues to forward the prune message to its upstream router.

Figure 46 RPT building at the multicast source side



As shown in Figure 46, the process for building a source-side RPT is relatively simple:

4. When a multicast source sends multicast packets to multicast group G, the DF in each network segment unconditionally forwards the packets to the RP.
5. The routers along the path from the source's directly connected router to the RP form an RPT branch. Each router on this branch adds a (*, G) entry to its forwarding table. The * means any multicast source.

After a bidirectional RPT is built, multicast traffic is forwarded along the source-side RPT and receiver-side RPT from sources to receivers.

NOTE:

If a receiver and a multicast source are at the same side of the RP, the source-side RPT and the receiver-side RPT might meet at a node before reaching the RP. In this case, multicast packets from the multicast source to the receiver are directly forwarded by the node to the receiver, instead of by the RP.

Administrative scoping overview

Division of PIM-SM domains

Typically, a PIM-SM domain or BIDIR-PIM domain contains only one BSR, which is responsible for advertising RP-set information within the entire PIM-SM/BIDIR-PIM domain. The information for all multicast groups is forwarded within the network scope that the BSR administers. This is called the "non-scoped BSR mechanism."

To implement refined management, you can divide a PIM-SM domain or BIDIR-PIM domain into one global scope zone and multiple administratively scoped zones (admin-scope zones). This is called the "administrative scoping mechanism."

The administrative scoping mechanism effectively releases stress on the management in a single-BSR domain and enables provision of zone-specific services through private group addresses.

Admin-scope zones are divided specific to multicast groups. Zone border routers (ZBRs) form the boundary of the admin-scope zone. Each admin-scope zone maintains one BSR, which serves multicast groups within a specific range. Multicast protocol packets, such as assert messages and bootstrap messages, for a specific group range cannot cross the admin-scope zone boundary. Multicast group ranges that different admin-scope zones serve can be overlapped. A multicast group is valid only within its local admin-scope zone, and functions as a private group address.

The global scope zone maintains a BSR, which serves the multicast groups that do not belong to any admin-scope zone.

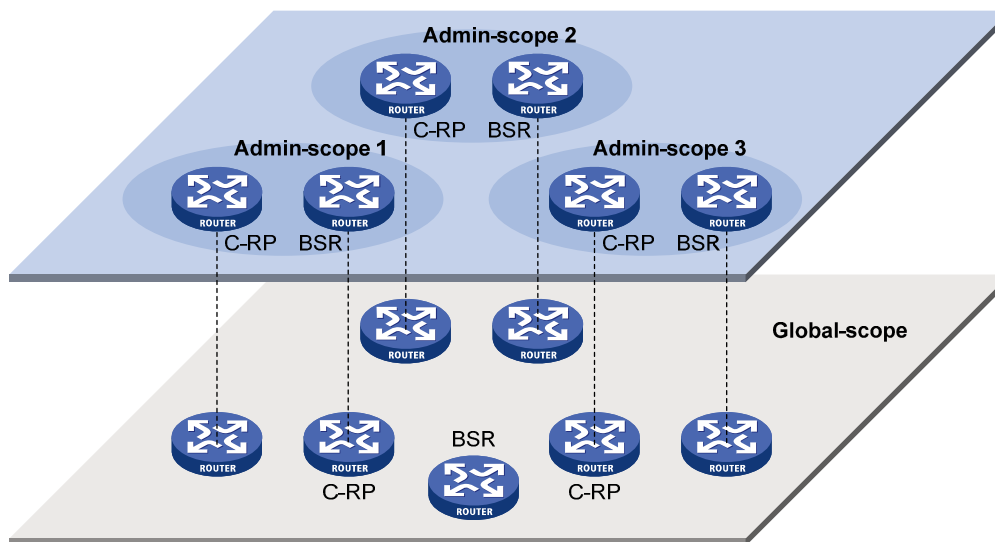
Relationship between admin-scope zones and the global scope zone

The global scope zone and each admin-scope zone have their own C-RPs and BSRs. These devices are effective only in their respective zones. That is, BSR election and RP election are implemented independently within each admin-scope zone. Each admin-scope zone has its own boundary. The multicast information cannot cross this boundary in either direction. A better understanding of the global scope zone and admin-scope zones should be based on geographical space and group address range.

1. Geographical space

Admin-scope zones are logical zones specific to particular multicast groups. The multicast packets of these multicast groups are confined within the local admin-scope zone and cannot cross the boundary of the zone.

Figure 47 Relationship between admin-scope zones and the global scope zone in geographic space

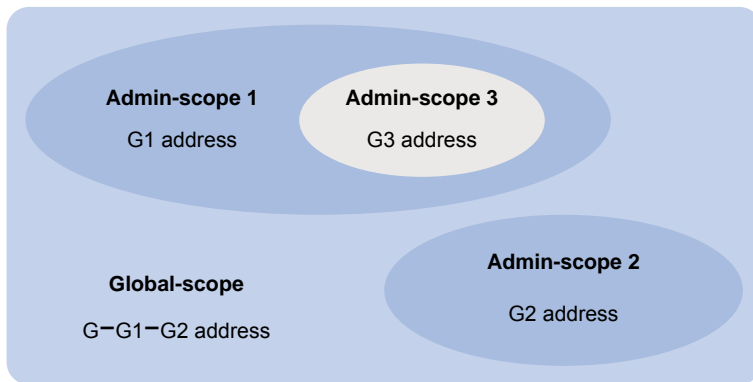


As shown in [Figure 47](#), for multicast groups in the same address range, admin-scope zones must be geographically separated from one another. Namely, a router must not serve different admin-scope zones. In other words, different admin-scope zones contain different routers, whereas the global scope zone covers all routers in the PIM-SM/BIDIR-PIM domain. Multicast packets that do not belong to any admin-scope zones can be transmitted in the entire PIM-SM/BIDIR-PIM domain.

2. Multicast group address ranges

Each admin-scope zone serves specific multicast groups. Usually, these addresses have no intersections. However, they might overlap one another.

Figure 48 Relationship between admin-scope zones and the global scope zone in group address ranges



In [Figure 48](#), the group address ranges of admin-scope 1 and 2 have no intersection, whereas the group address range of admin-scope 3 is a subset of the address range of admin-scope 1. The group address range of the global scope zone covers all the group addresses other than those of all the admin-scope zones. That is, the group address range of the global scope zone is G-G1-G2. In other words, a supplementary relationship exists between the global scope zone and all the admin-scope zones in terms of group address ranges.

PIM-SSM overview

The source-specific multicast (SSM) model and the any-source multicast (ASM) model are opposites. Presently, the ASM model includes the PIM-DM and PIM-SM modes. You can implement the SSM model by leveraging part of the PIM-SM technique. It is also called "PIM-SSM."

The SSM model provides a solution for source-specific multicast. It maintains the relationships between hosts and routers through IGMPv3.

In actual application, part of IGMPv3 or PIM-SM technique is adopted to implement the SSM model. In the SSM model, receivers locate a multicast source by means of advertisements, consultancy, and so on. No RP is needed, no RPT is required, no source registration process exists, and the multicast source discovery protocol (MSDP) is not needed for discovering sources in other PIM domains.

In PIM-SSM, the term "channel" refers to a multicast group, and the term "channel subscription" refers to a join message.

The working mechanism of PIM-SSM is summarized as follows:

- Neighbor discovery
- DR election
- SPT building

Neighbor discovery

PIM-SSM uses the same neighbor discovery mechanism as in PIM-DM and PIM-SM. See "[Neighbor discovery](#)."

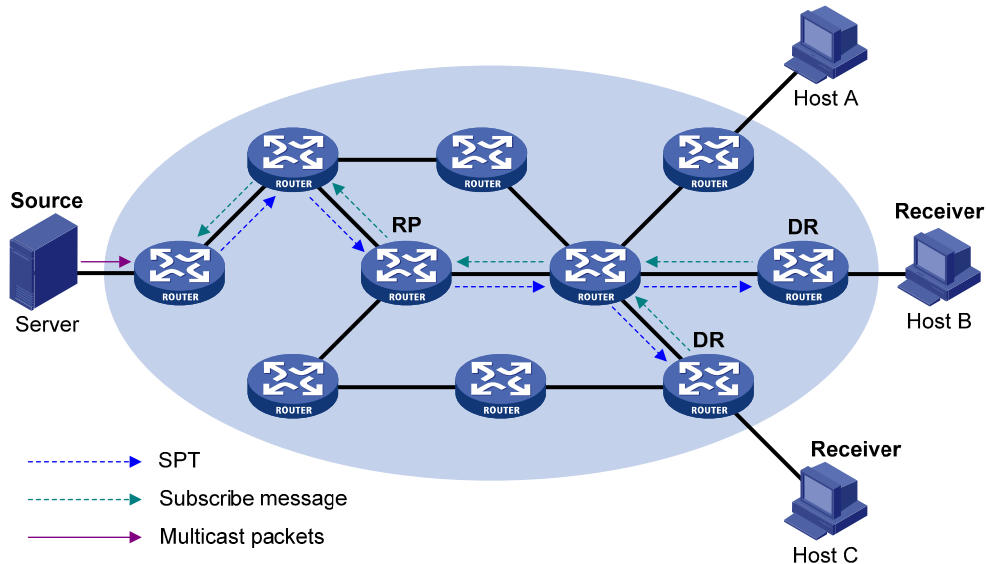
DR election

PIM-SSM uses the same DR election mechanism as in PIM-SM. See "[DR election](#)."

SPT building

The decision to build an RPT for PIM-SM or an SPT for PIM-SSM depends on whether the multicast group the receiver will join falls into the SSM group range (SSM group range reserved by IANA is 232.0.0.0/8).

Figure 49 SPT building in PIM-SSM



As shown in Figure 49, Host B and Host C are multicast information receivers. They send IGMPv3 report messages to the respective DRs to express their interest in the information about the specific multicast source S.

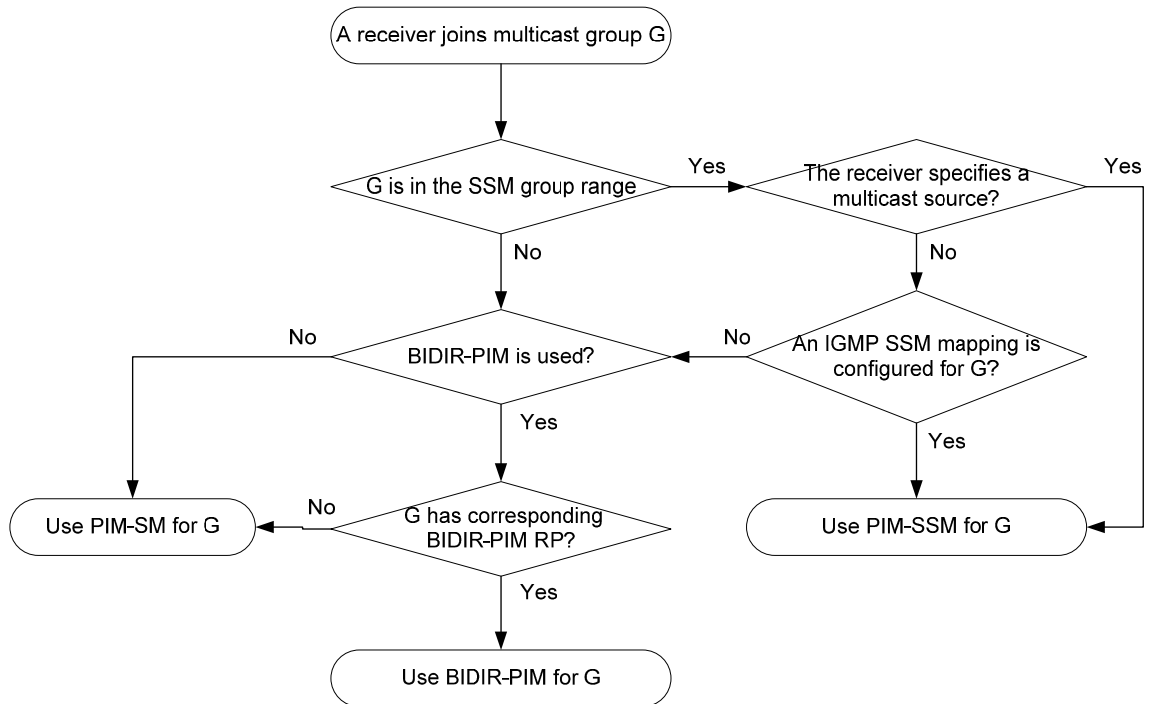
After receiving a report message, the DR first determines whether the group address in this message falls into the SSM group range and then does the following:

- If the group address in the message does fall into the SSM group range, the DR sends a subscribe message for channel subscription hop by hop toward the multicast source S. An (S, G) entry is created on all routers on the path from the DR to the source. An SPT is thereby built in the network, with the source S as its root and receivers as its leaves. This SPT is the transmission channel in PIM-SSM.
- If the group address in the message does not fall into the SSM group range, the receiver-side DR follows the PIM-SM process. The receiver-side DR sends a (*, G) join message to the RP, and the source-side DR registers the multicast source.

Relationships among PIM protocols

In a PIM network, PIM-DM cannot run together with PIM-SM, BIDIR-PIM, or PIM-SSM. However, PIM-SM, BIDIR-PIM, and PIM-SSM can run together. When they run together, which one is chosen for a receiver trying to join a group depends, as shown in Figure 50.

Figure 50 Relationships among PIM protocols



For more information about IGMP SSM mapping, see "[Configuring IGMP \(available only on the HP 5500 EI\)](#)."

PIM support for VPNs

To support PIM for VPNs, a multicast router that runs PIM maintains an independent set of PIM neighbor table, multicast routing table, BSR information, and RP-set information for each VPN.

After receiving a multicast data packet, the multicast router checks which VPN the data packet belongs to, and then forwards the packet according to the multicast routing table for that VPN or creates a multicast routing entry for that VPN.

Protocols and standards

- RFC 3973, *Protocol Independent Multicast-Dense Mode (PIM-DM): Protocol Specification(Revised)*
- RFC 4601, *Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification (Revised)*
- RFC 5015, *Bidirectional Protocol Independent Multicast (BIDIR-PIM)*
- RFC 5059, *Bootstrap Router (BSR) Mechanism for Protocol Independent Multicast (PIM)*
- RFC 4607, *Source-Specific Multicast for IP*
- Draft-ietf-ssm-overview-05, *An Overview of Source-Specific Multicast (SSM)*

Configuring PIM-DM

PIM-DM configuration task list

Task	Remarks
Enabling PIM-DM	Required
Enabling state-refresh capability	Optional
Configuring state-refresh parameters	Optional
Configuring PIM-DM graft retry period	Optional
Configuring PIM common features	Optional

Configuration prerequisites

Before you configure PIM-DM, complete the following tasks:

- Configure any unicast routing protocol so that all devices in the domain are interoperable at the network layer.
- Determine the interval between state-refresh messages.
- Determine the minimum time to wait before receiving a new refresh message.
- Determine the TTL value of state-refresh messages.
- Determine the graft retry period.

Enabling PIM-DM

With PIM-DM enabled, a router sends hello messages periodically to discover PIM neighbors and processes messages from the PIM neighbors. When you deploy a PIM-DM domain, enable PIM-DM on all non-border interfaces of the routers.

ⓘ IMPORTANT:

- All the interfaces in the same VPN instance on the same device must operate in the same PIM mode.
- PIM-DM does not work with multicast groups in the SSM group range.

Enabling PIM-DM globally on the public network

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IP multicast routing.	multicast routing-enable	Disabled by default.
3. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Enable PIM-DM.	pim dm	Disabled by default.

Enabling PIM-DM in a VPN instance

Step	Command	Description
1. Enter system view.	system-view	N/A
2. Create a VPN instance and enter VPN instance view.	ip vpn-instance <i>vpn-instance-name</i>	N/A

Step		Command	Description
3.	Configure an RD for the VPN instance.	route-distinguisher <i>route-distinguisher</i>	Not configured by default.
4.	Enable IP multicast routing.	multicast routing-enable	Disabled by default.
5.	Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
6.	Bind the interface with a VPN instance.	ip binding vpn-instance <i>vpn-instance-name</i>	By default, an interface belongs to the public network, and is not bound with any VPN instance.
7.	Enable PIM-DM.	pim dm	Disabled by default.

For more information about the **ip vpn-instance**, **route-distinguisher**, and **ip binding vpn-instance** commands, see *IP Routing Command Reference*.

For more information about the **multicast routing-enable** command, see *IP Multicast Command Reference*.

Enabling state-refresh capability

Pruned interfaces resume multicast forwarding when the pruned state times out. To prevent this, the router with the multicast source attached periodically sends an (S, G) state-refresh message, which is forwarded hop by hop along the initial multicast flooding path of the PIM-DM domain, to refresh the prune timer state of all the routers on the path. A multi-access subnet can have the state-refresh capability only if the state-refresh capability is enabled on all PIM routers on the subnet.

To enable the state-refresh capability:

Step		Command	Remarks
1.	Enter system view.	system-view	N/A
2.	Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3.	Enable the state-refresh capability.	pim state-refresh-capable	Optional Enabled by default

Configuring state-refresh parameters

The router directly connected with the multicast source periodically sends state-refresh messages. You can configure the interval for sending such messages.

A router might receive multiple state-refresh messages within a short time, and some of them might be duplicated messages. To keep a router from receiving such duplicated messages, you can configure the time that the router must wait before it receives next state-refresh message. If the router receives a new state-refresh message within the waiting time, it discards the message. If this timer times out, the router will accept a new state-refresh message, refresh its own PIM-DM state, and reset the waiting timer.

The TTL value of a state-refresh message decrements by 1 whenever it passes a router before it is forwarded to the downstream node until the TTL value comes down to 0. In a small network, a state-refresh message might cycle in the network. To effectively control the propagation scope of state-refresh messages, configure an appropriate TTL value based on the network size.

Perform the following configurations on all routers in the PIM domain.

To configure state-refresh parameters:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance vpn-instance-name]	N/A
3. Configure the interval between state-refresh messages.	state-refresh-interval interval	Optional 60 seconds by default
4. Configure the time to wait before receiving a new state-refresh message.	state-refresh-rate-limit interval	Optional 30 seconds by default
5. Configure the TTL value of state-refresh messages.	state-refresh-ttl ttl-value	Optional 255 by default

Configuring PIM-DM graft retry period

In PIM-DM, graft is the only type of message that uses the acknowledgment mechanism. In a PIM-DM domain, if a router does not receive a graft-ack message from the upstream router within the specified time after it sends a graft message, the router keeps sending new graft messages at a configurable interval (namely graft retry period), until it receives a graft-ack message from the upstream router.

To configure the graft retry period:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface interface-type interface-number	N/A
3. Configure the graft retry period.	pim timer graft-retry interval	Optional 3 seconds by default

For more information about the configuration of other timers in PIM-DM, see "[Configuring PIM common timers](#)."

Configuring PIM-SM

PIM-SM configuration task list

Task	Remarks	
Enabling PIM-SM	Required.	
Configuring an RP	Configuring a static RP	
	Configuring a C-RP	Required.
	Enabling auto-RP	Use any approach.

Task	Remarks	
	Configuring C-RP timers globally	Optional.
	Configuring a C-BSR	Required.
Configuring a BSR	Configuring a PIM domain border	Optional.
	Configuring global C-BSR parameters	Optional.
	Configuring C-BSR timers	Optional.
	Disabling BSM semantic fragmentation	Optional.
	Enabling administrative scoping	Optional.
Configuring administrative scoping	Configuring an admin-scope zone boundary	Optional.
	Configuring C-BSRs for each admin-scope zone and the global-scope zone	Optional.
Configuring multicast source registration	Optional.	
Disabling SPT switchover	Optional.	
Configuring PIM common features	Optional.	

Configuration prerequisites

Before you configure PIM-SM, complete the following tasks:

- Configure any unicast routing protocol so that all devices in the domain are interoperable at the network layer.
- Determine the IP address of a static RP and the ACL rule defining the range of multicast groups to be served by the static RP.
- Determine the C-RP priority and the ACL rule defining the range of multicast groups to be served by each C-RP.
- Determine the legal C-RP address range and the ACL rule defining the range of multicast groups to be served.
- Determine the C-RP-Adv interval.
- Determine the C-RP timeout.
- Determine the C-BSR priority.
- Determine the hash mask length.
- Determine the ACL rule defining a legal BSR address range.
- Determine the BS period.
- Determine the BS timeout.
- Determine the ACL rule for register message filtering.
- Determine the register suppression time.
- Determine the register probe time.
- Determine the ACL rule and sequencing rule for disabling an SPT switchover.

Enabling PIM-SM

With PIM-SM enabled, a router sends hello messages periodically to discover PIM neighbors and processes messages from the PIM neighbors. To deploy a PIM-SM domain, enable PIM-SM on all non-border interfaces of the routers.

! **IMPORTANT:**

All the interfaces in the same VPN instance on the same router must operate in the same PIM mode.

Enabling PIM-SM globally on the public network

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IP multicast routing.	multicast routing-enable	Disabled by default
3. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Enable PIM-SM.	pim sm	Disabled by default

Enabling PIM-SM in a VPN instance

Step	Command	Description
1. Enter system view.	system-view	N/A
2. Create a VPN instance and enter VPN instance view.	ip vpn-instance <i>vpn-instance-name</i>	N/A
3. Configure an RD for the VPN instance.	route-distinguisher <i>route-distinguisher</i>	Not configured by default.
4. Enable IP multicast routing.	multicast routing-enable	Disabled by default.
5. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
6. Bind the interface with a VPN instance.	ip binding vpn-instance <i>vpn-instance-name</i>	By default, an interface belongs to the public network, and is not bound with any VPN instance.
7. Enable PIM-SM.	pim sm	Disabled by default.

For more information about the **ip vpn-instance**, **route-distinguisher**, and **ip binding vpn-instance** commands, see *IP Routing Command Reference*.

For more information about the **multicast routing-enable** command, see *IP Multicast Command Reference*.

Configuring an RP

An RP can be manually configured or dynamically elected through the BSR mechanism. For a large PIM network, static RP configuration is a tedious job. Generally, static RP configuration is just a backup method for the dynamic RP election mechanism to enhance the robustness and operational manageability of a multicast network.

❗ **IMPORTANT:**

In a PIM network, if both PIM-SM and BIDIR-PIM are enabled, do not configure the same RP to serve PIM-SM and BIDIR-PIM simultaneously to avoid PIM routing table errors.

Configuring a static RP

If only one dynamic RP exists in a network, manually configuring a static RP can avoid communication interruption because of single-point failures. It can also avoid frequent message exchange between C-RPs and the BSR.

❗ **IMPORTANT:**

To enable a static RP to work normally, you must perform this configuration on all the routers in the PIM-SM domain and specify the same RP address.

Perform the following configuration on all the routers in the PIM-SM domain.

To configure a static RP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance vpn-instance-name]	N/A
3. Configure a static RP for PIM-SM.	static-rp rp-address [acl-number] [preferred]	No static RP by default

Configuring a C-RP

In a PIM-SM domain, you can configure routers that intend to become the RP as C-RPs. The BSR collects the C-RP information by receiving the C-RP-Adv messages from C-RPs or auto-RP announcements from other routers and organizes the information into an RP-set, which is flooded throughout the entire network. Then, the other routers in the network calculate the mappings between specific group ranges and the corresponding RPs based on the RP-set. HP recommends you to configure C-RPs on backbone routers.

To guard against C-RP spoofing, you must configure a legal C-RP address range and the range of multicast groups to be served on the BSR. In addition, because every C-BSR can become the BSR, you must configure the same filtering policy on all C-BSRs in the PIM-SM domain.

To configure a C-RP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance vpn-instance-name]	N/A
3. Configure an interface to be a C-RP for PIM-SM.	c-rp interface-type interface-number [group-policy acl-number priority priority holdtime hold-interval advertisement-interval adv-interval] *	No C-RPs are configured by default.
4. Configure a legal C-RP address range and the range of multicast groups to be served.	crp-policy acl-number	Optional. No restrictions by default.

NOTE:

- When you configure a C-RP, ensure a relatively large bandwidth between this C-RP and the other devices in the PIM-SM domain.
 - An RP can serve multiple multicast groups or all multicast groups. Only one RP can forward multicast traffic for a multicast group at a moment.
-

Enabling auto-RP

Auto-RP announcement and discovery messages are addressed to the multicast group addresses 224.0.1.39 and 224.0.1.40. With auto-RP enabled on a device, the device can receive these two types of messages and record the RP information carried in such messages.

To enable auto-RP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance vpn-instance-name]	N/A
3. Enable auto-RP.	auto-rp enable	Disabled by default

Configuring C-RP timers globally

To enable the BSR to distribute the RP-set information within the PIM-SM domain, C-RPs must periodically send C-RP-Adv messages to the BSR. The BSR learns the RP-set information from the received messages, and encapsulates its own IP address together with the RP-set information in its bootstrap messages. The BSR then floods the bootstrap messages to all PIM routers in the network.

Each C-RP encapsulates a timeout value in its C-RP-Adv messages. After receiving a C_RP-Adv message, the BSR obtains this timeout value and starts a C-RP timeout timer. If the BSR fails to hear a subsequent C-RP-Adv message from the C-RP when this timer times out, the BSR assumes the C-RP to have expired or become unreachable.

The C-RP timers need to be configured on C-RP routers.

To configure C-RP timers globally:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance vpn-instance-name]	N/A
3. Configure the C-RP-Adv interval.	c-rp advertisement-interval interval	Optional 60 seconds by default
4. Configure C-RP timeout time.	c-rp holdtime interval	Optional 150 seconds by default

For more information about the configuration of other timers in PIM-SM, see "[Configuring PIM common timers.](#)"

Configuring a BSR

Configuration guidelines

Before you configure a BSR, make sure that you are familiar with BSR election process, BSR legal address against BSR spoofing, PIM domain border, global C-BSR parameters, C-BSR timers, and bootstrap message fragments (BSMFs).

- BSR election process

A PIM-SM domain can have only one BSR, but must have at least one C-BSR. Any router can be configured as a C-BSR. Elected from C-BSRs, the BSR is responsible for collecting and advertising RP information in the PIM-SM domain.

C-BSRs should be configured on routers in the backbone network. When you configure a router as a C-BSR, be sure to specify a PIM-SM-enabled interface on the router. The BSR election process is summarized as follows:

- a. Initially, every C-BSR assumes itself to be the BSR of this PIM-SM domain and uses its interface IP address as the BSR address to send bootstrap messages.
- b. When a C-BSR receives the bootstrap message of another C-BSR, it first compares its own priority with the other C-BSR's priority carried in the message. The C-BSR with a higher priority wins. If a tie exists in the priority, the C-BSR with a higher IP address wins. The loser uses the winner's BSR address to replace its own BSR address and no longer assumes itself to be the BSR, and the winner retains its own BSR address and continues to assume itself to be the BSR.

- BSR legal address against BSR spoofing

Configuring a legal range of BSR addresses enables filtering of bootstrap messages based on the address range, thereby preventing a maliciously configured host from masquerading as a BSR. You must make the same configuration on all routers in the PIM-SM domain. The typical BSR spoofing cases and the corresponding preventive measures are as follows:

- Some maliciously configured hosts can forge bootstrap messages to fool routers and change RP mappings. Such attacks often occur on border routers. Because a BSR is inside the network whereas hosts are outside the network, you can protect a BSR against attacks from external hosts by enabling the border routers to perform neighbor checks and RPF checks on bootstrap messages and to discard unwanted messages.
- When an attacker controls a router in the network or when an illegal router is present in the network, the attacker can configure this router as a C-BSR and make it win BSR election to control the right of advertising RP information in the network. After a router is configured as a C-BSR, it automatically floods the network with bootstrap messages. Because a bootstrap message has a TTL value of 1, the whole network will not be affected as long as the neighbor router discards these bootstrap messages. Therefore, with a legal BSR address range configured on all routers in the entire network, all these routers will discard bootstrap messages from out of the legal address range.

These preventive measures can partially protect the security of BSRs in a network. However, if an attacker controls a legal BSR, the problem will still occur.

Because a large amount of information needs to be exchanged between a BSR and the other devices in the PIM-SM domain, a relatively large bandwidth should be provided between the C-BSRs and the other devices in the PIM-SM domain.

- PIM domain border

As the administrative core of a PIM-SM domain, the BSR sends the collected RP-set information in the form of bootstrap messages to all routers in the PIM-SM domain.

A PIM domain border is a bootstrap message boundary. Each BSR has its specific service scope. A number of PIM domain border interfaces partition a network into different PIM-SM domains. Bootstrap messages cannot cross a domain border in either direction.

- C-BSR parameters

In each PIM-SM domain, a unique BSR is elected from C-BSRs. The C-RPs in the PIM-SM domain send advertisement messages to the BSR. The BSR summarizes the advertisement messages to form an RP-set and advertises it to all routers in the PIM-SM domain. All the routers use the same hash algorithm to get the RP address that corresponds to specific multicast groups.

You can configure the hash mask length and C-BSR priority globally, in an admin-scope zone, and in the global scope zone.

- The values configured in the global scope zone or admin-scope zone have preference over the global values.
- If you do not configure these parameters in the global scope zone or admin-scope zone, the corresponding global values will be used.

For configuration of C-BSR parameters for an admin-scope zone and global scope zone, see "[Configuring C-BSRs for each admin-scope zone and the global-scope zone.](#)"

- C-BSR timers

The BSR election winner multicasts its own IP address and RP-set information through bootstrap messages within the entire zone it serves. The BSR floods bootstrap messages throughout the network at the interval of BS (BSR state) period. Any C-BSR that receives a bootstrap message retains the RP-set for the length of BS timeout, during which no BSR election takes place. If no bootstrap message is received from the BSR before the BS timeout timer expires, a new BSR election process is triggered among the C-BSRs.

About the BS period:

- By default, the BS period is determined by this formula: $BS\ period = (BS\ timeout - 10) / 2$. The default BS timeout is 130 seconds, so the default BS period = $(130 - 10) / 2 = 60$ (seconds).
- If this parameter is manually configured, the system will use the configured value.

About the BS timeout timer:

- By default, the BS timeout value is determined by this formula: $BS\ timeout\ timer = BS\ period \times 2 + 10$. The default BS period is 60 seconds, so the default BS timeout timer = $60 \times 2 + 10 = 130$ (seconds).
- If this parameter is manually configured, the system will use the configured value.

In the configuration, make sure that the BS period value is smaller than the BS timeout value.

- Bootstrap message fragments (BSMFs)

Generally, a BSR periodically distributes the RP-set information in bootstrap messages within the PIM-SM domain. It encapsulates a BSM in an IP datagram and might split the datagram into fragments if the message exceeds the maximum transmission unit (MTU). In respect of such IP fragmentation, loss of a single IP fragment leads to unavailability of the entire message.

Semantic fragmentation of BSMs can solve this issue. When a BSM exceeds the MTU, it is split to multiple bootstrap message fragments (BSMFs).

- After receiving a BSMF that contains the RP-set information of one group range, a non-BSR router updates corresponding RP-set information directly.
- If the RP-set information of one group range is carried in multiple BSMFs, a non-BSR router updates corresponding RP-set information after receiving all these BSMFs.

Because the RP-set information contained in each segment is different, loss of some IP fragments will not result in dropping of the entire message.

The function of BSM semantic fragmentation is enabled by default. Devices not supporting this function might deem a fragment as an entire message and learn only part of the RP-set information. Therefore, if such devices exist in the PIM-SM domain, you need to disable the semantic fragmentation function on the C-BSRs.

Generally, a BSR performs BSM semantic fragmentation according to the MTU of its BSR interface. However, the semantic fragmentation of BSMs originated because of learning of a new PIM neighbor is performed according to the MTU of the outgoing interface.

Configuring a C-BSR

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure an interface as a C-BSR.	c-bsr <i>interface-type</i> <i>interface-number</i> [<i>hash-length</i> [<i>priority</i>]]	No C-BSRs are configured by default.
4. Configure a legal BSR address range.	bsr-policy <i>acl-number</i>	Optional. No restrictions on BSR address range by default.

Configuring a PIM domain border

Perform the following configuration on routers that you want to configure as a PIM domain border.

To configure a PIM domain border:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure a PIM domain border.	pim bsr-boundary	By default, no PIM domain border is configured.

Configuring global C-BSR parameters

Perform the following configuration on C-BSR routers.

To configure C-BSR parameters:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure the hash mask length.	c-bsr hash-length <i>hash-length</i>	Optional. 30 by default.

Step	Command	Remarks
4. Configure the C-BSR priority.	c-bsr priority <i>priority</i>	Optional. By default, the C-BSR priority is 64.

Configuring C-BSR timers

Perform the following configuration on C-BSR routers.

To configure C-BSR timers:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure the BS period.	c-bsr interval <i>interval</i>	Optional. For the default value, see the note after this table.
4. Configure the BS timeout timer.	c-bsr holdtime <i>interval</i>	Optional. For the default value, see the note after this table.

Disabling BSM semantic fragmentation

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Disable the BSM semantic fragmentation function.	undo bsm-fragment enable	By default, the BSM semantic fragmentation function is enabled.

Configuring administrative scoping

When administrative scoping is disabled, a PIM-SM domain has only one BSR. The BSR manages the whole network. To manage your network more effectively and specifically, partition the PIM-SM domain into multiple admin-scope zones. Each admin-scope zone maintains a BSR, which serves a specific multicast group range. The global scope zone also maintains a BSR, which serves all the remaining multicast groups.

Enabling administrative scoping

Before you configure an admin-scope zone, you must enable administrative scoping.

Perform the following configuration on all routers in the PIM-SM domain.

To enable administrative scoping:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2.	Enter public network PIM view or VPN instance PIM view. pim [vpn-instance <i>vpn-instance-name</i>]	N/A
3.	Enable administrative scoping. c-bsr admin-scope	Disabled by default

Configuring an admin-scope zone boundary

ZBRs form the boundary of each admin-scope zone. Each admin-scope zone maintains a BSR, which serves a specific multicast group range. Multicast protocol packets (such as assert messages and bootstrap messages) that belong to this range cannot cross the admin-scope zone boundary.

Perform the following configuration on routers that you want to configure as a ZBR.

To configure an admin-scope zone boundary:

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Enter interface view. interface <i>interface-type</i> <i>interface-number</i>	N/A
3.	Configure a multicast forwarding boundary. multicast boundary <i>group-address</i> { <i>mask</i> <i>mask-length</i> }	By default, no multicast forwarding boundary is configured.

NOTE:

The *group-address* { *mask* | *mask-length* } parameter of the **multicast boundary** command can be used to specify the multicast groups an admin-scope zone serves, in the range of 239.0.0.0/8. For more information about the **multicast boundary** command, see *IP Multicast Command Reference*.

Configuring C-BSRs for each admin-scope zone and the global-scope zone

In a network with administrative scoping enabled, group-range-specific BSRs are elected from C-BSRs. C-RPs in the network send advertisement messages to the specific BSR. The BSR summarizes the advertisement messages to form an RP-set and advertises it to all routers in the specific admin-scope zone. All the routers use the same hash algorithm to get the RP address corresponding to the specific multicast group.

Configure C-BSRs for each admin-scope zone and the global-scope zone.

You can configure the hash mask length and C-BSR priority globally, in an admin-scope zone, and in the global scope zone.

- The values configured in the global scope zone or admin-scope zone have preference over the global values.
- If you do not configure these parameters in the global scope zone or admin-scope zone, the corresponding global values will be used.

For configuration of global C-BSR parameters, see "[Configuring global C-BSR parameters](#)."

- Configure C-BSRs for each admin-scope zone
Perform the following configuration on the routers that you want to configure as C-BSRs in admin-scope zones.
To configure a C-BSR for an admin-scope zone:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure a C-BSR for an admin-scope zone.	c-bsr group <i>group-address</i> { <i>mask</i> <i>mask-length</i> } [hash-length <i>hash-length</i> priority <i>priority</i>] *	No C-BSRs are configured for an admin-scope zone by default.

NOTE:

The *group-address* { *mask* | *mask-length* } parameter of the **c-bsr group** command can specify the multicast groups the C-BSR serves, in the range of 239.0.0.0/8.

- Configure C-BSRs for the global-scope zone
Perform the following configuration on the routers that you want to configure as C-BSRs in the global-scope zone.
To configure a C-BSR for the global-scope zone:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure a C-BSR for the global-scope zone.	c-bsr global [hash-length <i>hash-length</i> priority <i>priority</i>] *	No C-BSRs are configured for the global-scope zone by default.

Configuring multicast source registration

Within a PIM-SM domain, the source-side DR sends register messages to the RP, and these register messages have different multicast source or group addresses. You can configure a filtering rule to filter register messages so that the RP can serve specific multicast groups. If the filtering rule denies an (S, G) entry, or if the filtering rule does not define the action for this entry, the RP will send a register-stop message to the DR to stop the registration process for the multicast data.

Configuration guidelines

In view of information integrity of register messages in the transmission process, you can configure the device to calculate the checksum based on the entire register messages. However, to reduce the workload of encapsulating data in register messages and for the sake of interoperability, do not use this method of checksum calculation.

When receivers stop receiving multicast data addressed to a certain multicast group through the RP (that is, the RP stops serving the receivers of that multicast group), or when the RP starts receiving multicast data from the multicast source along the SPT, the RP sends a register-stop message to the source-side DR. After receiving this message, the DR stops sending register messages encapsulated with multicast data and starts a register-stop timer. Before the register-stop timer expires, the DR sends a null register message (a register message without encapsulated multicast data) to the RP. If the DR receives a register-stop message during the register probe time, it will reset its register-stop timer. Otherwise, the DR starts sending register messages with encapsulated data again when the register-stop timer expires.

The register-stop timer is set to a random value chosen uniformly from the interval (0.5 times register_suppression_time, 1.5 times register_suppression_time) minus register_probe_time.

Configure a filtering rule for register messages on all C-RP routers and configure them to calculate the checksum based on the entire register messages. Configure the register suppression time and the register probe time on all routers that might become source-side DRs.

Configuration procedure

To configure register-related parameters:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance vpn-instance-name]	N/A
3. Configure a filtering rule for register messages.	register-policy acl-number	Optional. No register filtering rule by default.
4. Configure the device to calculate the checksum based on the entire register messages.	register-whole-checksum	Optional. By default, the checksum is calculated based on the header of register messages.
5. Configure the register suppression time.	register-suppression-timeout interval	Optional. 60 seconds by default.
6. Configure the register probe time.	probe-interval interval	Optional 5 seconds by default.

Disabling SPT switchover

If an HP 5500 EI switch acts as an RP or the receiver-side DR, it initiates an SPT switchover process by default upon receiving the first multicast packet along the RPT. You can disable the switchover from RPT to SPT.

ⓘ IMPORTANT:

For an HP 5500 EI switch, once a multicast forwarding entry is created, subsequent multicast data will not be encapsulated in register messages before being forwarded even if a register outgoing interface is available. To avoid forwarding failure, do not use the **spt-switch-threshold infinity** command on a switch that might become an RP (a static RP or a C-RP).

To configure SPT switchover:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance vpn-instance-name]	N/A

Step	Command	Remarks
3. Disable SPT switchover.	<code>spt-switch-threshold infinity</code> [<code>group-policy acl-number</code> [<code>order order-value</code>]]	Optional. By default, the device switches to the SPT immediately after it receives the first multicast packet.

Configuring BIDIR-PIM

BIDIR-PIM configuration task list

Task	Remarks	
Enabling PIM-SM	Required.	
Enabling BIDIR-PIM	Required.	
Configuring an RP	Configuring a static RP	Required.
	Configuring a C-RP	Use any approach.
	Enabling auto-RP	
Configuring a BSR	Configuring C-RP timers globally	Optional.
	Configuring a C-BSR	Required.
	Configuring a BIDIR-PIM domain border	Optional.
	Configuring global C-BSR parameters	Optional.
	Configuring C-BSR timers	Optional.
Configuring an admin-scope zone boundary	Disabling BSM semantic fragmentation	Optional.
	Enabling administrative scoping	Optional.
	Configuring administrative scoping	Optional.
Configuring PIM common features	Configuring C-BSRs for each admin-scope zone and the global-scope zone	Optional.
		Optional.

Configuration prerequisites

Before you configure BIDIR-PIM, complete the following tasks:

- Configure a unicast routing protocol so that all devices in the domain can reach each other.
- Determine the IP address of a static RP and the ACL that defines the range of the multicast groups to be served by the static RP.
- Determine the C-RP priority and the ACL that defines the range of multicast groups to be served by each C-RP.
- Determine the legal C-RP address range and the ACL that defines the range of multicast groups to be served.
- Determine the C-RP-Adv interval.
- Determine the C-RP timeout.

- Determine the C-BSR priority.
- Determine the hash mask length.
- Determine the ACL defining the legal BSR address range.
- Determine the BS period.
- Determine the BS timeout.

Enabling PIM-SM

Because BIDIR-PIM is implemented on the basis of PIM-SM, you must enable PIM-SM before enabling BIDIR-PIM. To deploy a BIDIR-PIM domain, enable PIM-SM on all non-border interfaces of the domain.

ⓘ IMPORTANT:

On a router, all interfaces in the same VPN instance must operate in the same PIM mode.

Enabling PIM-SM globally for the public network

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IP multicast routing.	multicast routing-enable	Disabled by default
3. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Enable PIM-SM.	pim sm	Disabled by default

Enabling PIM-SM for a VPN instance

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a VPN instance and enter VPN instance view.	ip vpn-instance <i>vpn-instance-name</i>	N/A
3. Configure an RD for the VPN instance.	route-distinguisher <i>route-distinguisher</i>	Not configured by default.
4. Enable IP multicast routing.	multicast routing-enable	Disabled by default.
5. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
6. Bind the interface with the VPN instance.	ip binding vpn-instance <i>vpn-instance-name</i>	By default, an interface belongs to the public network, and is not bound with any VPN instance.
7. Enable PIM-SM.	pim sm	Disabled by default.

For more information about the **ip vpn-instance**, **route-distinguisher**, and **ip binding vpn-instance** commands, see *IP Routing Command Reference*.

For more information about the **multicast routing-enable** command, see *IP Multicast Command Reference*.

Enabling BIDIR-PIM

Perform this configuration on all routers in the BIDIR-PIM domain.

To enable BIDIR-PIM:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance vpn-instance-name]	N/A
3. Enable BIDIR-PIM.	bidir-pim enable	Disabled by default

Configuring an RP

An RP can be manually configured or dynamically elected through the BSR mechanism. For a large PIM network, static RP configuration is a tedious job. Generally, static RP configuration is just used as a backup means for the dynamic RP election mechanism to enhance the robustness and operation manageability of a multicast network.

⚠ CAUTION:

In a PIM network, if both PIM-SM and BIDIR-PIM are enabled, do not configure the same RP to serve PIM-SM and BIDIR-PIM simultaneously to avoid PIM routing table errors.

Configuring a static RP

If only one dynamic RP exists in a network, manually configuring a static RP can avoid communication interruption because of single-point failures and avoid frequent message exchange between C-RPs and the BSR.

In BIDIR-PIM, a static RP can be specified with a virtual IP address. For example, if the IP addresses of the interfaces at the two ends of a link are 10.1.1.1/24 and 10.1.1.2/24, you can specify a virtual IP address, like 10.1.1.100/24, for the static RP. As a result, the link becomes an RPL.

⚠ IMPORTANT:

You must perform static RP configuration on all routers in the PIM-SM domain and specify the same RP address.

Perform this configuration on all routers in the BIDIR-PIM domain.

To configure a static RP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance vpn-instance-name]	N/A
3. Configure a static RP for BIDIR-PIM.	static-rp rp-address [acl-number] [preferred] bidir	No static RP by default

Configuring a C-RP

In a BIDIR-PIM domain, you can configure routers that intend to become the RP as C-RPs. The BSR collects the C-RP information by receiving the C-RP-Adv messages from C-RPs or auto-RP announcements from other routers and organizes the information into an RP-set, which is flooded throughout the entire network. Then, the other routers in the network calculate the mappings between specific group ranges and the corresponding RPs based on the RP-set. HP recommends that you configure C-RPs on backbone routers.

To guard against C-RP spoofing, configure a legal C-RP address range and the range of multicast groups to be served on the BSR. In addition, because every C-BSR has a chance to become the BSR, you must configure the same filtering policy on all C-BSRs in the BIDIR-PIM domain.

To configure a C-RP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure an interface to be a C-RP for BIDIR-PIM.	c-rp <i>interface-type interface-number</i> [group-policy <i>acl-number</i> priority <i>priority</i> holdtime <i>hold-interval</i> advertisement-interval <i>adv-interval</i>] * bidir	No C-RP is configured by default.

NOTE:

- When you configure a C-RP, ensure a relatively large bandwidth between this C-RP and the other devices in the BIDIR-PIM domain.
- An RP can serve multiple multicast groups or all multicast groups. Only one RP can forward multicast traffic for a multicast group at a moment.

Enabling auto-RP

Auto-RP announcement and discovery messages are addressed to the multicast group addresses 224.0.1.39 and 224.0.1.40. With auto-RP enabled on a device, the device can receive these two types of messages and record the RP information carried in such messages.

To enable auto-RP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Enable auto-RP.	auto-rp enable	Disabled by default

Configuring C-RP timers globally

To enable the BSR to distribute the RP-set information within the BIDIR-PIM domain, C-RPs must periodically send C-RP-Adv messages to the BSR. The BSR learns the RP-set information from the received messages, and encapsulates its own IP address together with the RP-set information in its bootstrap messages. The BSR then floods the bootstrap messages to all PIM routers in the network.

Each C-RP encapsulates a timeout value in its C-RP-Adv messages. After receiving a C_RP-Adv message, the BSR obtains this timeout value and starts a C-RP timeout timer. If the BSR fails to hear a subsequent C-RP-Adv message from the C-RP within the timeout interval, the BSR assumes the C-RP to have expired or become unreachable.

The C-RP timers need to be configured on C-RP routers.

To configure C-RP timers globally:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance vpn-instance-name]	N/A
3. Configure the C-RP-Adv interval.	c-rp advertisement-interval interval	Optional 60 seconds by default
4. Configure C-RP timeout time.	c-rp holdtime interval	Optional 150 seconds by default

For more information about the configuration of other timers in BIDIR-PIM, see "[Configuring PIM common timers.](#)"

Configuring a BSR

Configuration guidelines

Before you configure a BSR, make sure that you are familiar with BSR election process, BSR legal address against BSR spoofing, PIM domain border, global C-BSR parameters, C-BSR timers, and bootstrap message fragments (BSMFs). They operate the same way as in the PIM-SM domain. For more information, see "[Configuration guidelines.](#)"

Configuring a C-BSR

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance vpn-instance-name]	N/A
3. Configure an interface as a C-BSR.	c-bsr interface-type interface-number [hash-length [priority]]	No C-BSRs are configured by default.
4. Configure a legal BSR address range.	bsr-policy acl-number	Optional. No restrictions on BSR address range by default.

Configuring a BIDIR-PIM domain border

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure a BIDIR-PIM domain border.	pim bsr-boundary	By default, no BIDIR-PIM domain border is configured.

Configuring global C-BSR parameters

Perform the following configuration on C-BSR routers.

To configure global C-BSR parameters:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure the hash mask length.	c-bsr hash-length <i>hash-length</i>	Optional. 30 by default.
4. Configure the C-BSR priority.	c-bsr priority <i>priority</i>	Optional. 64 by default.

Configuring C-BSR timers

Perform the following configuration on C-BSR routers.

To configure C-BSR timers:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure the BS period.	c-bsr interval <i>interval</i>	Optional. For the default value, see the note after this table.
4. Configure the BS timeout timer.	c-bsr holdtime <i>interval</i>	Optional. For the default value, see the note after this table.

Disabling BSM semantic fragmentation

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Disable the BSM semantic fragmentation function.	undo bsm-fragment enable	By default, the BSM semantic fragmentation function is enabled.

Configuring administrative scoping

With administrative scoping disabled, a BIDIR-PIM domain has only one BSR. The BSR manages the whole network. To manage your network more effectively and specifically, partition the BIDIR-PIM domain into multiple admin-scope zones. Each admin-scope zone maintains a BSR, which serves a specific multicast group range. The global scope zone also maintains a BSR, which serves all the rest multicast groups.

Enabling administrative scoping

Before you configure an admin-scope zone, you must enable administrative scoping first.

Perform the following configuration on all routers in the BIDIR-PIM domain.

To enable administrative scoping:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance vpn-instance-name]	N/A
3. Enable administrative scoping.	c-bsr admin-scope	Disabled by default

Configuring an admin-scope zone boundary

The boundary of each admin-scope zone is formed by ZBRs. Each admin-scope zone maintains a BSR, which serves a specific multicast group range. Multicast protocol packets (such as assert messages and bootstrap messages) that belong to this range cannot cross the admin-scope zone boundary.

Perform the following configuration on routers that you want to configure as a ZBR.

To configure an admin-scope zone boundary:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure a multicast forwarding boundary.	multicast boundary <i>group-address</i> { <i>mask</i> <i>mask-length</i> }	By default, no multicast forwarding boundary is configured.

NOTE:

The *group-address* { *mask* | *mask-length* } part of the **multicast boundary** command can be used to specify the multicast groups an admin-scope zone serves, in the range of 239.0.0.0/8. For more information about the **multicast boundary** command, see *IP Multicast Command Reference*.

Configuring C-BSRs for each admin-scope zone and the global-scope zone

In a network with administrative scoping enabled, group-range-specific BSRs are elected from C-BSRs. C-RPs in the network send advertisement messages to the specific BSR. The BSR summarizes the advertisement messages to form an RP-set and advertises it to all routers in the specific admin-scope zone.

All the routers use the same hash algorithm to get the RP address corresponding to the specific multicast group.

Configure C-BSRs for each admin-scope zone and the global-scope zone.

You can configure the hash mask length and C-BSR priority globally, in an admin-scope zone, and in the global scope zone.

- The values configured in the global scope zone or admin-scope zone have preference over the global values.
- If you do not configure these parameters in the global scope zone or admin-scope zone, the corresponding global values will be used.

For configuration of global C-BSR parameters, see "[Configuring global C-BSR parameters.](#)"

- Configure C-BSRs for each admin-scope zone
Perform the following configuration on the routers that you want to configure as C-BSRs in admin-scope zones.

To configure a C-BSR for an admin-scope zone:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure a C-BSR for an admin-scope zone.	c-bsr group <i>group-address</i> { <i>mask</i> <i>mask-length</i> } [hash-length <i>hash-length</i> priority <i>priority</i>] *	No C-BSRs are configured for an admin-scope zone by default.

NOTE:

The *group-address* { *mask* | *mask-length* } parameter of the **c-bsr group** command can specify the multicast groups the C-BSR serves, in the range of 239.0.0.0/8.

- Configure C-BSRs for the global-scope zone
Perform the following configuration on the routers that you want to configure as C-BSRs in the global-scope zone.
To configure a C-BSR for the global-scope zone:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure a C-BSR for the global-scope zone.	c-bsr global [hash-length <i>hash-length</i> priority <i>priority</i>] *	No C-BSRs are configured for the global-scope zone by default.

Configuring PIM-SSM

! IMPORTANT:

PIM-SSM needs the support of IGMPv3. Be sure to enable IGMPv3 on PIM routers with multicast receivers.

PIM-SSM configuration task list

Task	Remarks
Enabling PIM-SM	Required
Configuring the SSM group range	Optional
Configuring PIM common features	Optional

Configuration prerequisites

Before you configure PIM-SSM, complete the following tasks:

- Configure any unicast routing protocol so that all devices in the domain are interoperable at the network layer.
- Determine the SSM group range.

Enabling PIM-SM

The implementation of the SSM model is based on some subsets of PIM-SM. Therefore, you must enable PIM-SM before configuring PIM-SSM.

When deploying a PIM-SSM domain, enable PIM-SM on non-border interfaces of the routers.

 **IMPORTANT:**

All the interfaces in the same VPN instance on the same device must operate in the same PIM mode.

Enabling PIM-SM globally on the public network

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IP multicast routing.	multicast routing-enable	Disabled by default
3. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Enable PIM-SM.	pim sm	Disabled by default

Enabling PIM-SM in a VPN instance

Step	Command	Description
1. Enter system view.	system-view	N/A
2. Create a VPN instance and enter VPN instance view.	ip vpn-instance <i>vpn-instance-name</i>	N/A
3. Configure an RD for the VPN instance.	route-distinguisher <i>route-distinguisher</i>	No RD is configured by default.
4. Enable IP multicast routing.	multicast routing-enable	Disabled by default.

Step	Command	Description
5. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
6. Bind the interface with a VPN instance.	ip binding vpn-instance <i>vpn-instance-name</i>	By default, an interface belongs to the public network, and it is not bound with any VPN instance.
7. Enable PIM-SM.	pim sm	Disabled by default.

For more information about the **ip vpn-instance**, **route-distinguisher**, and **ip binding vpn-instance** commands, see *IP Routing Command Reference*.

For more information about the **multicast routing-enable** command, see *IP Multicast Command Reference*.

Configuring the SSM group range

Configuration guidelines

As for whether the information from a multicast source is delivered to the receivers based on the PIM-SSM model or the PIM-SM model, this depends on whether the group address in the (S, G) channel subscribed by the receivers falls into the SSM group range. All PIM-SM-enabled interfaces assume that multicast groups within this address range are using the PIM-SSM model.

Make sure that the same SSM group range is configured on all routers in the entire domain. Otherwise, multicast information cannot be delivered through the SSM model.

When a member of a multicast group in the SSM group range sends an IGMPv1 or IGMPv2 report message, the device does not trigger a (*, G) join.

Configuration procedure

Perform the following configuration on all routers in the PIM-SSM domain.

To configure an SSM multicast group range:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure the SSM group range.	ssm-policy <i>acl-number</i>	Optional 232.0.0.0/8 by default

Configuring PIM common features

For the functions or parameters that can be configured in both PIM view and interface view described in this section:

- In PIM view, the configuration is effective on all interfaces. In interface view, the configuration is effective on only the current interface.

- If the same function or parameter is configured in both PIM view and interface view, the configuration in interface view has preference over the configuration in PIM view, regardless of the configuration sequence.

PIM common feature configuration task list

Task	Remarks
Configuring a multicast data filter	Optional
Configuring a hello message filter	Optional
Configuring PIM hello options	Optional
Configuring the prune delay	Optional
Configuring PIM common timers	Optional
Configuring join/prune message sizes	Optional
Configuring PIM to work with BFD	Optional
Setting the DSCP value for PIM messages	Optional

Configuration prerequisites

Before you configure PIM common features, complete the following tasks:

- Configure any unicast routing protocol so that all devices in the domain are interoperable at the network layer.
- Configure PIM-DM, or PIM-SM, or PIM-SSM.
- Determine the ACL rule for filtering multicast data.
- Determine the ACL rule defining a legal source address range for hello messages.
- Determine the priority for DR election (global value/interface level value).
- Determine the PIM neighbor timeout time (global value/interface value).
- Determine the prune message delay (global value/interface level value).
- Determine the prune override interval (global value/interface level value).
- Determine the prune delay.
- Determine the hello interval (global value/interface level value).
- Determine the maximum delay between hello message (interface level value).
- Determine the assert timeout time (global value/interface value).
- Determine the join/prune interval (global value/interface level value).
- Determine the join/prune timeout (global value/interface value).
- Determine the multicast source lifetime.
- Determine the maximum size of join/prune messages.
- Determine the maximum number of (S, G) entries in a join/prune message.
- Determine the DSCP value for PIM messages.

Configuring a multicast data filter

In either a PIM-DM domain or a PIM-SM domain, routers can check passing-by multicast data based on the configured filtering rules and determine whether to continue forwarding the multicast data. In other words, PIM routers can act as multicast data filters. These filters can help implement traffic control on one hand, and control the information available to receivers downstream to enhance data security on the other hand.

Generally, a smaller distance from the filter to the multicast source results in a more remarkable filtering effect.

This filter works not only on independent multicast data but also on multicast data encapsulated in register messages.

To configure a multicast data filter:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure a multicast group filter.	source-policy <i>acl-number</i>	No multicast data filter by default

Configuring a hello message filter

Along with the wide applications of PIM, the security requirement for the protocol is becoming increasingly demanding. The establishment of correct PIM neighboring relationships is the prerequisite for secure application of PIM. You can configure a legal source address range for hello messages on interfaces of routers to ensure the correct PIM neighboring relationships, guarding against PIM message attacks.

To configure a hello message filter:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure a hello message filter.	pim neighbor-policy <i>acl-number</i>	No hello message filter is configured by default.

NOTE:

With the hello message filter configured, if hello messages of an existing PIM neighbor fail to pass the filter, the PIM neighbor will be removed automatically when it times out.

Configuring PIM hello options

PIM hello options

In either a PIM-DM domain or a PIM-SM domain, the hello messages sent among routers contain the following configurable options:

- **DR_Priority** (for PIM-SM only)—Priority for DR election. The device with the highest priority wins the DR election. You can configure this parameter on all the routers in a multi-access network directly connected to multicast sources or receivers.
- **Holdtime**—The timeout time of PIM neighbor reachability state. When this timer times out, if the router has received no hello message from a neighbor, it assumes that this neighbor has expired or become unreachable.
- **LAN_Prune_Delay**—The delay of prune messages on a multi-access network. This option consists of LAN-delay (namely, prune message delay), override-interval, and neighbor tracking flag. If the LAN-delay or override-interval values of different PIM routers on a multi-access subnet are different, the largest value takes effect. If you want to enable neighbor tracking, be sure to enable the neighbor tracking feature on all PIM routers on a multi-access subnet.

The LAN-delay setting will cause the upstream routers to delay processing received prune messages. The override-interval sets the length of time that a downstream router can wait before sending a prune override message. When a router receives a prune message from a downstream router, it does not perform the prune action immediately. Instead, it maintains the current forwarding state for a period of LAN-delay plus override-interval. If the downstream router needs to continue receiving multicast data, it must send a join message within the prune override interval. Otherwise, the upstream router will perform the prune action when the period of LAN-delay plus override-interval times out.

A hello message sent from a PIM router contains a generation ID option. The generation ID is a random value for the interface on which the hello message is sent. Normally, the generation ID of a PIM router does not change unless the status of the router changes (for example, when PIM is just enabled on the interface or the device is restarted). When the router starts or restarts sending hello messages, it generates a new generation ID. If a PIM router finds that the generation ID in a hello message from the upstream router has changed, it assumes that the status of the upstream neighbor is lost or that the upstream neighbor has changed. In this case, it triggers a join message for state update.

If you disable join suppression (namely, enable neighbor tracking), be sure to disable the join suppression feature on all PIM routers on a multi-access subnet. Otherwise, the upstream router will fail to explicitly track join messages from downstream routers.

Configuring hello options globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance vpn-instance-name]	N/A
3. Configure the priority for DR election.	hello-option dr-priority priority	Optional. 1 by default.
4. Configure PIM neighbor timeout time.	hello-option holdtime interval	Optional. 105 seconds by default.
5. Configure the prune message delay time (LAN-delay).	hello-option lan-delay interval	Optional. 500 milliseconds by default.
6. Configure the prune override interval.	hello-option override-interval interval	Optional. 2500 milliseconds by default.
7. Disable join suppression.	hello-option neighbor-tracking	Enabled by default.

Configuring hello options on an interface

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the priority for DR election.	pim hello-option dr-priority <i>priority</i>	Optional. 1 by default.
4. Configure PIM neighbor timeout time.	pim hello-option holdtime <i>interval</i>	Optional. 105 seconds by default.
5. Configure the prune message delay time (LAN-delay).	pim hello-option lan-delay <i>interval</i>	Optional. 500 milliseconds by default.
6. Configure the prune override interval.	pim hello-option override-interval <i>interval</i>	Optional. 2,00 milliseconds by default.
7. Disable join suppression.	pim hello-option neighbor-tracking	Enabled by default.
8. Configure the interface to reject hello messages without a generation ID.	pim require-genid	By default, hello messages without Generation_ID are accepted.

Configuring the prune delay

Configuring a prune delay interval on an upstream router on a shared network segment can make the upstream router not perform the prune action immediately after it receives the prune message from its downstream router. Instead, the upstream router maintains the current forwarding state for a period of time that the prune delay interval defines. In this period, if the upstream router receives a join message from the downstream router, it cancels the prune action. Otherwise, it performs the prune action.

To configure the prune delay time:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure the prune delay interval.	prune delay <i>interval</i>	Optional. 3 seconds by default, which equals the prune pending time.

Configuring PIM common timers

PIM common timers

PIM routers discover PIM neighbors and maintain PIM neighboring relationships with other routers by periodically sending out hello messages.

After receiving a hello message, a PIM router waits a random period, which is smaller than the maximum delay between hello messages, before sending a hello message. This delay avoids collisions that occur when multiple PIM routers send hello messages simultaneously.

A PIM router periodically sends join/prune messages to its upstream for state update. A join/prune message contains the join/prune timeout time. The upstream router sets a join/prune timeout timer for each pruned downstream interface.

Any router that has lost assert election will prune its downstream interface and maintain the assert state for a period of time. When the assert state times out, the assert losers will resume multicast forwarding.

When a router fails to receive subsequent multicast data from multicast source S, the router does not immediately delete the corresponding (S, G) entry. Instead, it maintains the (S, G) entry for a period of time (namely, the multicast source lifetime) before deleting the (S, G) entry.

Configuring PIM common timers globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance vpn-instance-name]	N/A
3. Configure the hello interval.	timer hello interval	Optional. 30 seconds by default.
4. Configure the join/prune interval.	timer join-prune interval	Optional. 60 seconds by default.
5. Configure the join/prune timeout time.	holdtime join-prune interval	Optional. 210 seconds by default.
6. Configure assert timeout time.	holdtime assert interval	Optional. 180 seconds by default.
7. Configure the multicast source lifetime.	source-lifetime interval	Optional. 210 seconds by default.

Configuring PIM common timers on an interface

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface interface-type interface-number	N/A
3. Configure the hello interval.	pim timer hello interval	Optional. 30 seconds by default.
4. Configure the maximum delay between hello messages.	pim triggered-hello-delay interval	Optional. 5 seconds by default.
5. Configure the join/prune interval.	pim timer join-prune interval	Optional. 60 seconds by default.

Step	Command	Remarks
6. Configure the join/prune timeout time.	pim holdtime join-prune interval	Optional. 210 seconds by default.
7. Configure assert timeout time.	pim holdtime assert interval	Optional. 180 seconds by default.

NOTE:

If no special networking requirements are raised, use the default settings.

Configuring join/prune message sizes

A large size of a join/prune message might result in loss of a larger amount of information if a message is lost. You can set a small value for the size of each join/prune message to reduce the impact in case of the loss of a message.

By controlling the maximum number of (S, G) entries in a join/prune message, you can effectively reduce the number of (S, G) entries sent per unit of time.

! **IMPORTANT:**

If PIM snooping-enabled switches are deployed in the PIM network, be sure to set a value no greater than the path MTU for the maximum size of each join/prune message on the receiver-side edge PIM devices

To configure join/prune message sizes:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network PIM view or VPN instance PIM view.	pim [vpn-instance vpn-instance-name]	N/A
3. Configure the maximum size of each join/prune message.	ip-pkt-size packet-size	Optional. 8100 bytes by default.
4. Configure the maximum number of (S, G) entries in a join/prune message.	ip-queue-size queue-size	Optional. 1020 by default.

Configuring PIM to work with BFD

PIM uses hello messages to elect a DR for a multi-access network. The elected DR will be the only multicast forwarder on the multi-access network.

If the DR fails, a new DR election process will start after the DR is aged out. However, it might take a long period of time. To start a new DR election process immediately after the original DR fails, enable PIM to work with Bidirectional Forwarding Detection (BFD) on a multi-access network to detect failures of the links among PIM neighbors. You must enable PIM to work with BFD on all PIM-capable routers on a multi-access network, so that the PIM neighbors can fast detect DR failures and start a new DR election process.

For more information about BFD, see *High Availability Configuration Guide*.

Before you configure this feature on an interface, be sure to enable PIM-DM or PIM-SM on the interface. To enable PIM to work with BFD:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type interface-number</i>	N/A
3. Enable PIM to work with BFD.	pim bfd enable	Disabled by default

Setting the DSCP value for PIM messages

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network IGMP view or VPN instance PIM view.	pim [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Set the DSCP value for PIM messages	dscp <i>dscp-value</i>	Optional. By default, the DSCP value in PIM messages is 48.

Displaying and maintaining PIM

Task	Command	Remarks
Display the BSR information in the PIM-SM domain and locally configured C-RP information in effect.	display pim [all-instance vpn-instance <i>vpn-instance-name</i>] bsr-info [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the information of unicast routes used by PIM.	display pim [all-instance vpn-instance <i>vpn-instance-name</i>] claimed-route [<i>source-address</i>] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the number of PIM control messages.	display pim [all-instance vpn-instance <i>vpn-instance-name</i>] control-message counters [message-type { probe register register-stop }] [interface <i>interface-type interface-number</i> message-type { assert bsr crp graft graft-ack hello join-prune state-refresh }] *] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display the DF information of BIDIR-PIM.	display pim [all-instance vpn-instance <i>vpn-instance-name</i>] df-info [<i>rp-address</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the information about unacknowledged PIM-DM graft messages.	display pim [all-instance vpn-instance <i>vpn-instance-name</i>] grafts [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the PIM information on an interface or all interfaces.	display pim [all-instance vpn-instance <i>vpn-instance-name</i>] interface [<i>interface-type</i> <i>interface-number</i>] [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the information of join/prune messages to send.	display pim [all-instance vpn-instance <i>vpn-instance-name</i>] join-prune mode { sm [flags <i>flag-value</i>] ssm } [interface <i>interface-type</i> <i>interface-number</i> neighbor <i>neighbor-address</i>] * [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display PIM neighboring information.	display pim [all-instance vpn-instance <i>vpn-instance-name</i>] neighbor [interface <i>interface-type</i> <i>interface-number</i> <i>neighbor-address</i> verbose] * [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the content of the PIM routing table.	display pim [all-instance vpn-instance <i>vpn-instance-name</i>] routing-table [<i>group-address</i> [mask { <i>mask-length</i> <i>mask</i> }] <i>source-address</i> [mask { <i>mask-length</i> <i>mask</i> }]] incoming-interface [<i>interface-type</i> <i>interface-number</i> register] outgoing-interface { include exclude match } { <i>interface-type</i> <i>interface-number</i> register } mode <i>mode-type</i> flags <i>flag-value</i> fsm] * [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the RP information.	display pim [all-instance vpn-instance <i>vpn-instance-name</i>] rp-info [<i>group-address</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Reset PIM control message counters.	reset pim [all-instance vpn-instance <i>vpn-instance-name</i>] control-message counters [interface <i>interface-type</i> <i>interface-number</i>]	Available in user view

PIM configuration examples

PIM-DM configuration example

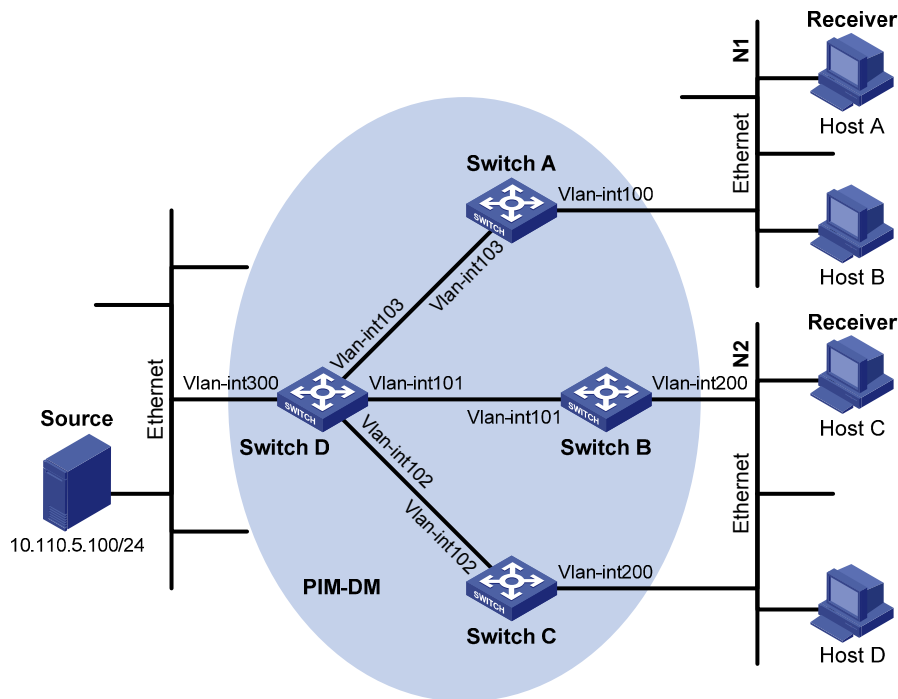
Network requirements

As shown in Figure 51, receivers receive VOD information through multicast. The receiver groups of different organizations form stub networks, and one or more receiver hosts exist in each stub network. The entire PIM domain is operating in the dense mode.

Host A and Host C are multicast receivers in two stub networks.

IGMPv2 runs between Switch A and N1 and between Switch B/Switch C and N2.

Figure 51 Network diagram



Device	Interface	IP address	Device	Interface	IP address
Switch A	Vlan-int100	10.110.1.1/24	Switch D	Vlan-int300	10.110.5.1/24
	Vlan-int103	192.168.1.1/24		Vlan-int103	192.168.1.2/24
Switch B	Vlan-int200	10.110.2.1/24		Vlan-int101	192.168.2.2/24
	Vlan-int101	192.168.2.1/24		Vlan-int102	192.168.3.2/24
Switch C	Vlan-int200	10.110.2.2/24			
	Vlan-int102	192.168.3.1/24			

Configuration procedure

1. Configure IP addresses and unicast routing:

Configure the IP address and subnet mask for each interface as per [Figure 51](#). (Details not shown.)
Configure OSPF on the switches in the PIM-DM domain to ensure network-layer reachability among them. (Details not shown.)

2. Enable IP multicast routing, and enable PIM-DM and IGMP:

Enable IP multicast routing on Switch A, enable PIM-DM on each interface, and enable IGMP on VLAN-interface 100, which connects Switch A to the stub network.

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] igmp enable
[SwitchA-Vlan-interface100] pim dm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 103
[SwitchA-Vlan-interface103] pim dm
[SwitchA-Vlan-interface103] quit
```

The configuration on Switch B and Switch C is similar to that on Switch A.

Enable IP multicast routing on Switch D, and enable PIM-DM on each interface.

```
<SwitchD> system-view
[SwitchD] multicast routing-enable
[SwitchD] interface vlan-interface 300
[SwitchD-Vlan-interface300] pim dm
[SwitchD-Vlan-interface300] quit
[SwitchD] interface vlan-interface 103
[SwitchD-Vlan-interface103] pim dm
[SwitchD-Vlan-interface103] quit
[SwitchD] interface vlan-interface 101
[SwitchD-Vlan-interface101] pim dm
[SwitchD-Vlan-interface101] quit
[SwitchD] interface vlan-interface 102
[SwitchD-Vlan-interface102] pim dm
[SwitchD-Vlan-interface102] quit
```

3. Verify the configuration:

Display PIM configuration information on Switch D.

```
[SwitchD] display pim interface
VPN-Instance: public net

```

Interface	NbrCnt	HelloInt	DR-Pri	DR-Address	
Vlan300	0	30	1	10.110.5.1	(local)
Vlan103	1	30	1	192.168.1.2	(local)
Vlan101	1	30	1	192.168.2.2	(local)
Vlan102	1	30	1	192.168.3.2	(local)

Display PIM neighboring relationships on Switch D.

```
[SwitchD] display pim neighbor
VPN-Instance: public net
Total Number of Neighbors = 3
```

Neighbor	Interface	Uptime	Expires	Dr-Priority
192.168.1.1	Vlan103	00:02:22	00:01:27	1
192.168.2.1	Vlan101	00:00:22	00:01:29	3
192.168.3.1	Vlan102	00:00:23	00:01:31	5

Assume that Host A needs to receive the information addressed to multicast group G (225.1.1.1). After multicast source S (10.110.5.100/24) sends multicast packets to the multicast group G, an SPT is established through traffic flooding. Switches on the SPT path (Switch A and Switch D) have their (S, G) entries. Host A sends an IGMP report to Switch A to join the multicast group G, and a (*, G) entry is generated on Switch A. You can use the **display pim routing-table** command to view the PIM routing table information on each switch. For example:

Display PIM routing table information on Switch A.

```
[SwitchA] display pim routing-table
VPN-Instance: public net
Total 1 (*, G) entry; 1 (S, G) entry

(*, 225.1.1.1)
  Protocol: pim-dm, Flag: WC
  UpTime: 00:04:25
  Upstream interface: NULL
    Upstream neighbor: NULL
    RPF prime neighbor: NULL
  Downstream interface(s) information:
  Total number of downstreams: 1
    1: Vlan-interface100
      Protocol: igmp, UpTime: 00:04:25, Expires: never

(10.110.5.100, 225.1.1.1)
  Protocol: pim-dm, Flag: ACT
  UpTime: 00:06:14
  Upstream interface: Vlan-interface103
    Upstream neighbor: 192.168.1.2
    RPF prime neighbor: 192.168.1.2
  Downstream interface(s) information:
  Total number of downstreams: 1
    1: Vlan-interface100
      Protocol: pim-dm, UpTime: 00:04:25, Expires: never
```

Display PIM routing table information on Switch D.

```
[SwitchD] display pim routing-table
VPN-Instance: public net
Total 0 (*, G) entry; 1 (S, G) entry

(10.110.5.100, 225.1.1.1)
  Protocol: pim-dm, Flag: LOC ACT
  UpTime: 00:03:27
  Upstream interface: Vlan-interface300
    Upstream neighbor: NULL
    RPF prime neighbor: NULL
```

```

Downstream interface(s) information:
Total number of downstreams: 2
  1: Vlan-interface103
      Protocol: pim-dm, UpTime: 00:03:27, Expires: never
  2: Vlan-interface102
      Protocol: pim-dm, UpTime: 00:03:27, Expires: never

```

PIM-SM non-scoped zone configuration example

Network requirements

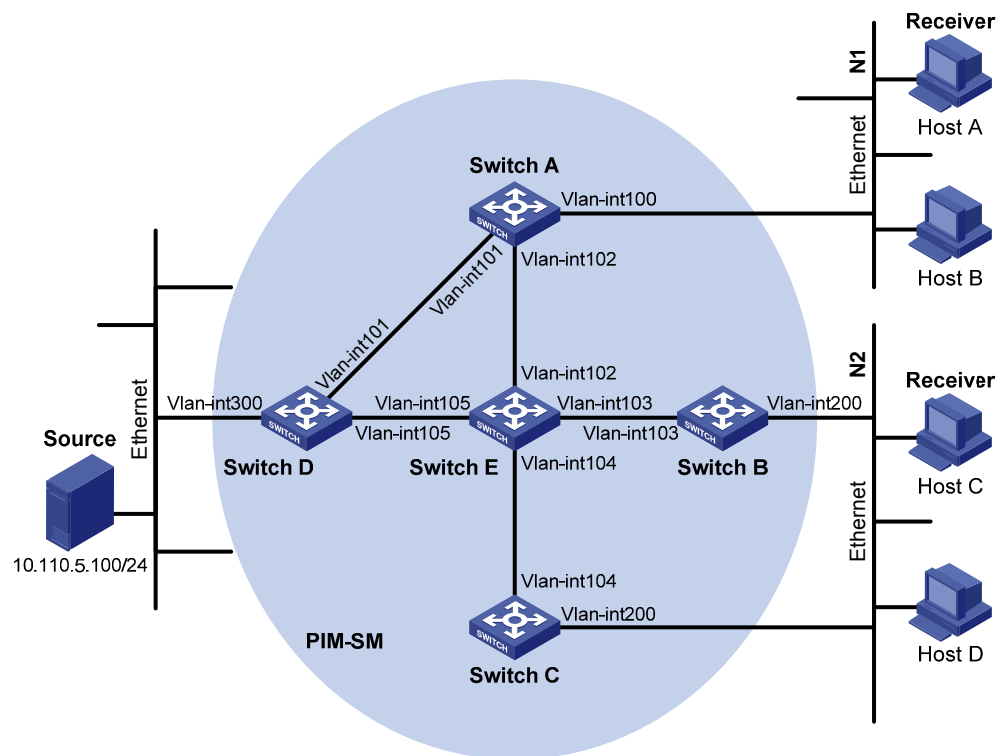
As shown in [Figure 52](#), receivers receive VOD information through multicast. The receiver groups of different organizations form stub networks, and one or more receiver hosts exist in each stub network. The entire PIM-SM domain contains only one BSR.

Host A and Host C are multicast receivers in two stub networks.

VLAN-interface 105 on Switch D and VLAN-interface 102 on Switch E act as C-BSRs and C-RPs. The C-BSR on Switch E has a higher priority. The multicast group range served by the C-RP is 225.1.1.0/24. Modify the hash mask length to map a certain number of consecutive group addresses within the range to the two C-RPs.

IGMPv2 runs between Switch A and N1 and between Switch B/Switch C and N2.

Figure 52 Network diagram



Device	Interface	IP address	Device	Interface	IP address
Switch A	Vlan-int100	10.110.1.1/24	Switch D	Vlan-int300	10.110.5.1/24
	Vlan-int101	192.168.1.1/24		Vlan-int101	192.168.1.2/24
	Vlan-int102	192.168.9.1/24		Vlan-int105	192.168.4.2/24
Switch B	Vlan-int200	10.110.2.1/24	Switch E	Vlan-int104	192.168.3.2/24

	Vlan-int103	192.168.2.1/24	Vlan-int103	192.168.2.2/24
Switch C	Vlan-int200	10.110.2.2/24	Vlan-int102	192.168.9.2/24
	Vlan-int104	192.168.3.1/24	Vlan-int105	192.168.4.1/24

Configuration procedure

1. Configure IP addresses and unicast routing:

Configure the IP address and subnet mask for each interface as per [Figure 52](#). (Details not shown.)

Configure OSPF on the switches in the PIM-SM domain to ensure network-layer reachability among them. (Details not shown.)

2. Enable IP multicast routing, and enable PIM-SM and IGMP:

Enable IP multicast routing on Switch A, enable PIM-SM on each interface, and enable IGMP on VLAN-interface 100, which connects Switch A to the stub network.

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] igmp enable
[SwitchA-Vlan-interface100] pim sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim sm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 102
[SwitchA-Vlan-interface102] pim sm
[SwitchA-Vlan-interface102] quit
```

The configuration on Switch B and Switch C is similar to that on Switch A. The configuration on Switch D and Switch E is also similar to that on Switch A except that it is not necessary to enable IGMP on the corresponding interfaces on these two switches.

3. Configure a C-BSR and a C-RP:

On Switch D, configure the service scope of RP, specify a C-BSR and a C-RP, and set the hash mask length to 32 and the priority of the C-BSR to 10.

```
<SwitchD> system-view
[SwitchD] acl number 2005
[SwitchD-acl-basic-2005] rule permit source 225.1.1.0 0.0.0.255
[SwitchD-acl-basic-2005] quit
[SwitchD] pim
[SwitchD-pim] c-bsr vlan-interface 105 32 10
[SwitchD-pim] c-rp vlan-interface 105 group-policy 2005
[SwitchD-pim] quit
```

On Switch E, configure the service scope of RP advertisements, specify a C-BSR and a C-RP, and set the hash mask length to 32 and the priority of the C-BSR to 20.

```
<SwitchE> system-view
[SwitchE] acl number 2005
[SwitchE-acl-basic-2005] rule permit source 225.1.1.0 0.0.0.255
[SwitchE-acl-basic-2005] quit
[SwitchE] pim
[SwitchE-pim] c-bsr vlan-interface 102 32 20
```

```
[SwitchE-pim] c-rp vlan-interface 102 group-policy 2005
[SwitchE-pim] quit
```

4. Verify the configuration:

Display PIM configuration information on Switch A.

```
[SwitchA] display pim interface
VPN-Instance: public net
Interface          NbrCnt HelloInt  DR-Pri  DR-Address
Vlan100            0       30       1       10.110.1.1 (local)
Vlan101            1       30       1       192.168.1.2
Vlan102            1       30       1       192.168.9.2
```

Display BSR information and the locally configured C-RP information in effect on Switch A.

```
[SwitchA] display pim bsr-info
VPN-Instance: public net
Elected BSR Address: 192.168.9.2
  Priority: 20
  Hash mask length: 32
  State: Accept Preferred
  Scope: Not scoped
  Uptime: 00:40:40
  Expires: 00:01:42
```

Display BSR information and the locally configured C-RP information in effect on Switch D.

```
[SwitchD] display pim bsr-info
VPN-Instance: public net
Elected BSR Address: 192.168.9.2
  Priority: 20
  Hash mask length: 32
  State: Accept Preferred
  Scope: Not scoped
  Uptime: 00:05:26
  Expires: 00:01:45
Candidate BSR Address: 192.168.4.2
  Priority: 10
  Hash mask length: 32
  State: Candidate
  Scope: Not scoped

Candidate RP: 192.168.4.2(Vlan-interface105)
  Priority: 192
  HoldTime: 150
  Advertisement Interval: 60
  Next advertisement scheduled at: 00:00:34
```

Display BSR information and the locally configured C-RP information in effect on Switch E.

```
[SwitchE] display pim bsr-info
VPN-Instance: public net
Elected BSR Address: 192.168.9.2
  Priority: 20
  Hash mask length: 32
```

```
State: Elected
Scope: Not scoped
Uptime: 00:01:18
Next BSR message scheduled at: 00:01:52
Candidate BSR Address: 192.168.9.2
Priority: 20
Hash mask length: 32
State: Elected
Scope: Not scoped
```

```
Candidate RP: 192.168.9.2(Vlan-interface102)
Priority: 192
HoldTime: 150
Advertisement Interval: 60
Next advertisement scheduled at: 00:00:48
```

Display RP information on Switch A.

```
[SwitchA] display pim rp-info
```

```
VPN-Instance: public net
PIM-SM BSR RP information:
Group/MaskLen: 225.1.1.0/24
RP: 192.168.4.2
Priority: 192
HoldTime: 150
Uptime: 00:51:45
Expires: 00:02:22
```

```
RP: 192.168.9.2
Priority: 192
HoldTime: 150
Uptime: 00:51:45
Expires: 00:02:22
```

Assume that Host A needs to receive information addressed to the multicast group G (225.1.1.0). The RP corresponding to the multicast group G is Switch E as a result of hash calculation, so an RPT will be built between Switch A and Switch E. When the multicast source S (10.110.5.100/24) registers with the RP, an SPT will be built between Switch D and Switch E. After receiving multicast data, Switch A immediately switches from the RPT to the SPT. Switches on the RPT path (Switch A and Switch E) have a (*, G) entry, and switches on the SPT path (Switch A and Switch D) have an (S, G) entry. You can use the **display pim routing-table** command to view the PIM routing table information on the switches. For example:

Display PIM routing table information on Switch A.

```
[SwitchA] display pim routing-table
```

```
VPN-Instance: public net
Total 1 (*, G) entry; 1 (S, G) entry
```

```
(*, 225.1.1.0)
RP: 192.168.9.2
Protocol: pim-sm, Flag: WC
UpTime: 00:13:46
```

```
Upstream interface: Vlan-interface102
  Upstream neighbor: 192.168.9.2
  RPF prime neighbor: 192.168.9.2
Downstream interface(s) information:
Total number of downstreams: 1
  1: Vlan-interface100
    Protocol: igmp, UpTime: 00:13:46, Expires: 00:03:06
```

```
(10.110.5.100, 225.1.1.0)
RP: 192.168.9.2
Protocol: pim-sm, Flag: SPT ACT
UpTime: 00:00:42
Upstream interface: Vlan-interface101
  Upstream neighbor: 192.168.1.2
  RPF prime neighbor: 192.168.1.2
Downstream interface(s) information:
Total number of downstreams: 1
  1: Vlan-interface100
    Protocol: pim-sm, UpTime: 00:00:42, Expires: 00:03:06
```

The information on Switch B and Switch C is similar to that on Switch A.

Display PIM routing table information on Switch D.

```
[SwitchD] display pim routing-table
VPN-Instance: public net
Total 0 (*, G) entry; 1 (S, G) entry
```

```
(10.110.5.100, 225.1.1.0)
RP: 192.168.9.2
Protocol: pim-sm, Flag: SPT LOC ACT
UpTime: 00:00:42
Upstream interface: Vlan-interface300
  Upstream neighbor: NULL
  RPF prime neighbor: NULL
Downstream interface(s) information:
Total number of downstreams: 1
  1: Vlan-interface105
    Protocol: pim-sm, UpTime: 00:00:42, Expires: 00:02:26
```

Display PIM routing table information on Switch E.

```
[SwitchE] display pim routing-table
VPN-Instance: public net
Total 1 (*, G) entry; 0 (S, G) entry
```

```
(*, 225.1.1.0)
RP: 192.168.9.2 (local)
Protocol: pim-sm, Flag: WC
UpTime: 00:13:16
Upstream interface: Register
  Upstream neighbor: 192.168.4.2
  RPF prime neighbor: 192.168.4.2
```

```
Downstream interface(s) information:
Total number of downstreams: 1
  1: Vlan-interface102
      Protocol: pim-sm, UpTime: 00:13:16, Expires: 00:03:22
```

PIM-SM admin-scope zone configuration example

Network requirements

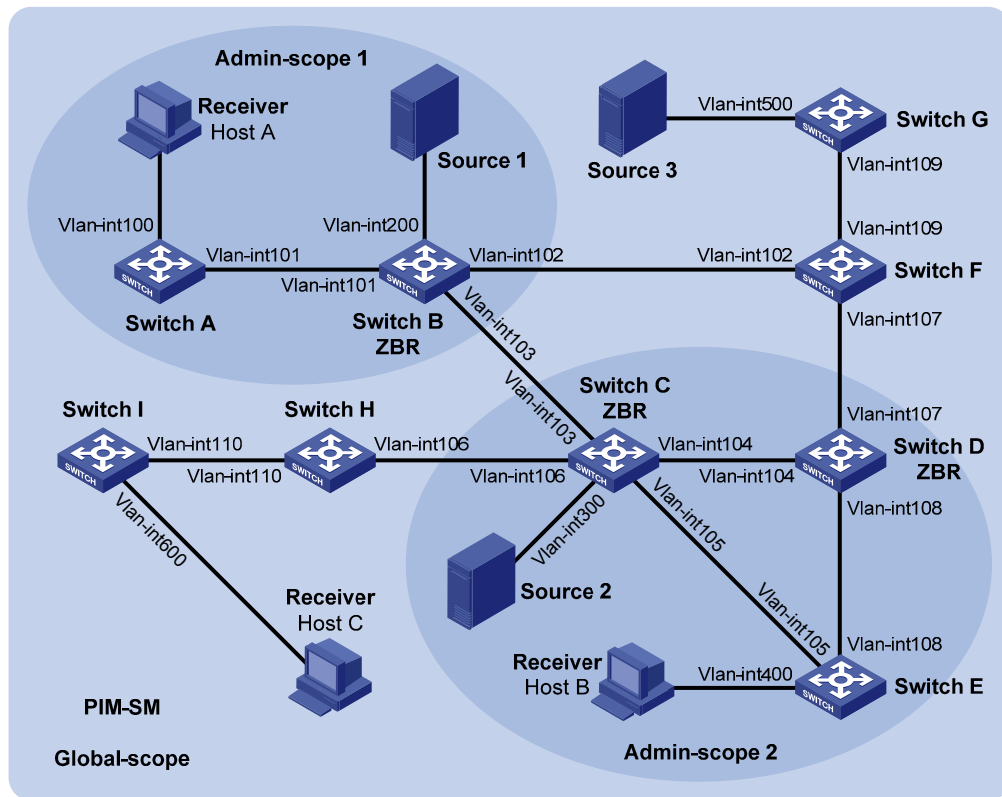
As shown in [Figure 53](#), receivers receive VOD information through multicast. The entire PIM-SM domain is divided into admin-scope zone 1, admin-scope zone 2, and the global zone. Switch B, Switch C, and Switch D are ZBRs of these three domains respectively.

Source 1 and Source 2 send different multicast information to multicast group 239.1.1.1. Host A receives the multicast information only from Source 1, and Host B receives the multicast information only from Source 2. Source 3 sends multicast information to multicast group 224.1.1.1. Host C is a multicast receiver for this multicast group.

VLAN-interface 101 of Switch B acts as a C-BSR and C-RP of admin-scope zone 1, which serves the multicast group range 239.0.0.0/8. VLAN-interface 104 of Switch D acts as a C-BSR and C-RP of admin-scope zone 2, which also serves the multicast group range 239.0.0.0/8. VLAN-interface 109 of Switch F acts as C-BSRs and C-RPs of the global scope zone, which serves all the multicast groups other than those in the 239.0.0.0/8 range.

IGMPv2 runs between Switch A, Switch E, Switch I, and their respective receivers.

Figure 53 Network diagram



Device	Interface	IP address	Device	Interface	IP address
Switch A	Vlan-int100	192.168.1.1/24	Switch D	Vlan-int104	10.110.4.2/24
	Vlan-int101	10.110.1.1/24		Vlan-int108	10.110.7.1/24
Switch B	Vlan-int200	192.168.2.1/24	Switch E	Vlan-int400	192.168.4.1/24
	Vlan-int101	10.110.1.2/24		Vlan-int105	10.110.5.2/24
	Vlan-int103	10.110.2.1/24		Vlan-int108	10.110.7.2/24
	Vlan-int102	10.110.3.1/24	Switch F	Vlan-int109	10.110.9.1/24
Switch C	Vlan-int300	192.168.3.1/24		Vlan-int107	10.110.8.2/24
	Vlan-int104	10.110.4.1/24		Vlan-int102	10.110.3.2/24
	Vlan-int105	10.110.5.1/24	Switch G	Vlan-int500	192.168.5.1/24
	Vlan-int103	10.110.2.2/24		Vlan-int109	10.110.9.2/24
	Vlan-int106	10.110.6.1/24	Source 1	—	192.168.2.10/24
Switch H	Vlan-int110	10.110.10.1/24	Source 2	—	192.168.3.10/24
	Vlan-int106	10.110.6.2/24	Source 3	—	192.168.5.10/24
Switch I	Vlan-int600	192.168.6.1/24			
	Vlan-int110	10.110.10.2/24			

Configuration procedure

1. Configure IP addresses and unicast routing:
 Configure the IP address and subnet mask for each interface as per Figure 53. (Details not shown.)
 Configure OSPF on the switches in the PIM-SM domain to ensure network-layer reachability among them. (Details not shown.)
2. Enable IP multicast routing and administrative scoping, and enable PIM-SM and IGMP:

Enable IP multicast routing and administrative scoping on Switch A, enable PIM-SM on each interface, and enable IGMP on the host-side interface VLAN-interface 100.

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] pim
[SwitchA-pim] c-bsr admin-scope
[SwitchA-pim] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] igmp enable
[SwitchA-Vlan-interface100] pim sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim sm
[SwitchA-Vlan-interface101] quit
```

The configuration on Switch E and Switch I is similar to the configuration on Switch A.

On Switch B, enable IP multicast routing and administrative scoping, and enable PIM-SM on each interface.

```
<SwitchB> system-view
[SwitchB] multicast routing-enable
[SwitchB] pim
[SwitchB-pim] c-bsr admin-scope
[SwitchB-pim] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] pim sm
[SwitchB-Vlan-interface200] quit
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] pim sm
[SwitchB-Vlan-interface101] quit
[SwitchB] interface vlan-interface 102
[SwitchB-Vlan-interface102] pim sm
[SwitchB-Vlan-interface102] quit
[SwitchB] interface vlan-interface 103
[SwitchB-Vlan-interface103] pim sm
[SwitchB-Vlan-interface103] quit
```

The configuration on Switch C, Switch D, Switch F, Switch G, and Switch H is similar to the configuration on Switch B. (Details not shown.)

3. Configure an admin-scope zone boundary:

On Switch B, configure VLAN-interface 102 and VLAN-interface 103 to be the boundary of admin-scope zone 1.

```
[SwitchB] interface vlan-interface 102
[SwitchB-Vlan-interface102] multicast boundary 239.0.0.0 8
[SwitchB-Vlan-interface102] quit
[SwitchB] interface vlan-interface 103
[SwitchB-Vlan-interface103] multicast boundary 239.0.0.0 8
[SwitchB-Vlan-interface103] quit
```

On Switch C, configure VLAN-interface 103 and VLAN-interface 106 to be the boundary of admin-scope zone 2.

```

<SwitchC> system-view
[SwitchC] interface vlan-interface 103
[SwitchC-Vlan-interface103] multicast boundary 239.0.0.0 8
[SwitchC-Vlan-interface103] quit
[SwitchC] interface vlan-interface 106
[SwitchC-Vlan-interface106] multicast boundary 239.0.0.0 8
[SwitchC-Vlan-interface106] quit

```

On Switch D, configure VLAN-interface 107 to be the boundary of admin-scope zone 2.

```

<SwitchD> system-view
[SwitchD] interface vlan-interface 107
[SwitchD-Vlan-interface107] multicast boundary 239.0.0.0 8
[SwitchD-Vlan-interface107] quit

```

4. Configure C-BSRs and C-RPs:

On Switch B, configure the service scope of RP advertisements, and configure VLAN-interface 101 as a C-BSR and C-RP of admin-scope zone 1.

```

[SwitchB] acl number 2001
[SwitchB-acl-basic-2001] rule permit source 239.0.0.0 0.255.255.255
[SwitchB-acl-basic-2001] quit
[SwitchB] pim
[SwitchB-pim] c-bsr group 239.0.0.0 8
[SwitchB-pim] c-bsr vlan-interface 101
[SwitchB-pim] c-rp vlan-interface 101 group-policy 2001
[SwitchB-pim] quit

```

On Switch D, configure the service scope of RP advertisements, and configure VLAN-interface 104 as a C-BSR and C-RP of admin-scope zone 2.

```

[SwitchD] acl number 2001
[SwitchD-acl-basic-2001] rule permit source 239.0.0.0 0.255.255.255
[SwitchD-acl-basic-2001] quit
[SwitchD] pim
[SwitchD-pim] c-bsr group 239.0.0.0 8
[SwitchD-pim] c-bsr vlan-interface 104
[SwitchD-pim] c-rp vlan-interface 104 group-policy 2001
[SwitchD-pim] quit

```

On Switch F, configure VLAN-interface 109 as a C-BSR and C-RP in the global scope zone.

```

<SwitchF> system-view
[SwitchF] pim
[SwitchF-pim] c-bsr global
[SwitchF-pim] c-bsr vlan-interface 109
[SwitchF-pim] c-rp vlan-interface 109
[SwitchF-pim] quit

```

5. Verify the configuration:

Display BSR information and the locally configured C-RP information on Switch B.

```

[SwitchB] display pim bsr-info
VPN-Instance: public net
Elected BSR Address: 10.110.9.1
  Priority: 64
  Hash mask length: 30

```



```
State: Accept Preferred
Scope: Global
Uptime: 00:01:45
Expires: 00:01:25
Elected BSR Address: 10.110.1.2
Priority: 64
Hash mask length: 30
State: Elected
Scope: 239.0.0.0/8
Uptime: 00:04:54
Next BSR message scheduled at: 00:00:06
Candidate BSR Address: 10.110.1.2
Priority: 64
Hash mask length: 30
State: Elected
Scope: 239.0.0.0/8
```

```
Candidate RP: 10.110.1.2(Vlan-interface101)
Priority: 192
HoldTime: 150
Advertisement Interval: 60
Next advertisement scheduled at: 00:00:15
```

Display BSR information and the locally configured C-RP information on Switch D.

```
[SwitchD] display pim bsr-info
```

```
VPN-Instance: public net
Elected BSR Address: 10.110.9.1
Priority: 64
Hash mask length: 30
State: Accept Preferred
Scope: Global
Uptime: 00:01:45
Expires: 00:01:25
Elected BSR Address: 10.110.4.2
Priority: 64
Hash mask length: 30
State: Elected
Scope: 239.0.0.0/8
Uptime: 00:03:48
Next BSR message scheduled at: 00:01:12
Candidate BSR Address: 10.110.4.2
Priority: 64
Hash mask length: 30
State: Elected
Scope: 239.0.0.0/8
```

```
Candidate RP: 10.110.4.2(Vlan-interface104)
Priority: 192
HoldTime: 150
```

```
Advertisement Interval: 60
Next advertisement scheduled at: 00:00:10
```

Display BSR information and the locally configured C-RP information on Switch F.

```
[SwitchF] display pim bsr-info
VPN-Instance: public net
Elected BSR Address: 10.110.9.1
  Priority: 64
  Hash mask length: 30
  State: Elected
  Scope: Global
  Uptime: 00:11:11
  Next BSR message scheduled at: 00:00:49
Candidate BSR Address: 10.110.9.1
  Priority: 64
  Hash mask length: 30
  State: Elected
  Scope: Global
```

```
Candidate RP: 10.110.9.1(Vlan-interface109)
  Priority: 192
  HoldTime: 150
  Advertisement Interval: 60
  Next advertisement scheduled at: 00:00:55
```

Display RP information on Switch B.

```
[SwitchB] display pim rp-info
VPN-Instance: public net
PIM-SM BSR RP information:
Group/MaskLen: 224.0.0.0/4
  RP: 10.110.9.1
  Priority: 192
  HoldTime: 150
  Uptime: 00:03:39
  Expires: 00:01:51
```

```
Group/MaskLen: 239.0.0.0/8
  RP: 10.110.1.2 (local)
  Priority: 192
  HoldTime: 150
  Uptime: 00:07:44
  Expires: 00:01:51
```

Display RP information on Switch D.

```
[SwitchD] display pim rp-info
VPN-Instance: public net
PIM-SM BSR RP information:
Group/MaskLen: 224.0.0.0/4
  RP: 10.110.9.1
  Priority: 192
  HoldTime: 150
```

```

Uptime: 00:03:42
Expires: 00:01:48

```

```

Group/MaskLen: 239.0.0.0/8
RP: 10.110.4.2 (local)
Priority: 192
HoldTime: 150
Uptime: 00:06:54
Expires: 00:02:41

```

Display RP information on Switch F.

```

[SwitchF] display pim rp-info
VPN-Instance: public net
PIM-SM BSR RP information:
Group/MaskLen: 224.0.0.0/4
RP: 10.110.9.1 (local)
Priority: 192
HoldTime: 150
Uptime: 00:00:32
Expires: 00:01:58

```

BIDIR-PIM configuration example

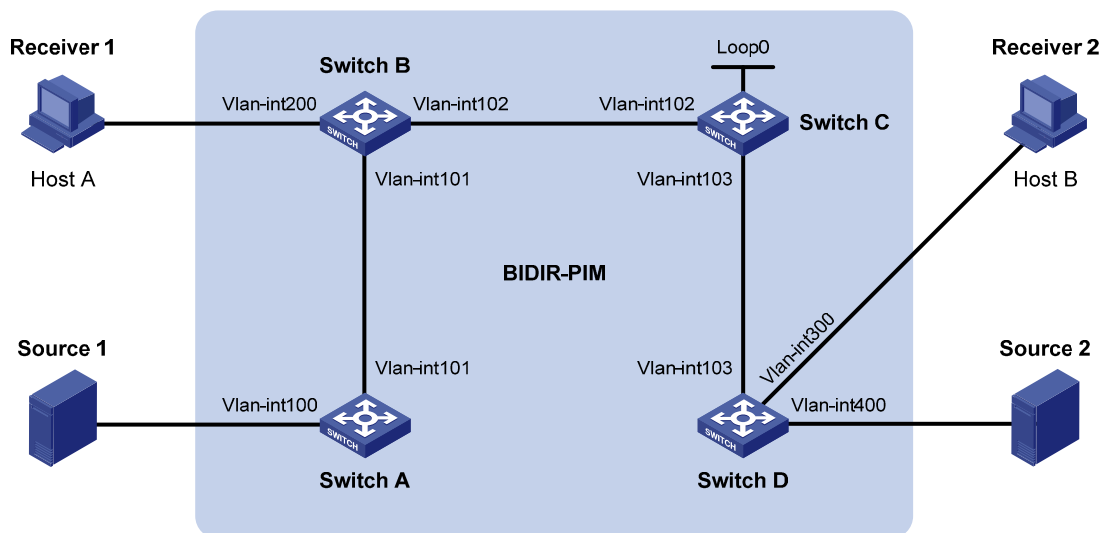
Network requirements

In the BIDIR-PIM domain shown in [Figure 54](#). Source 1 and Source 2 send different multicast information to multicast group 225.1.1.1. Host A and Host B receive multicast information from the two sources.

VLAN interface 102 of Switch C acts as a C-BSR, and loopback interface 0 of Switch C acts as a C-RP of the BIDIR-PIM domain.

IGMPv2 runs between Switch B and Host A, and between Switch D and Host B.

Figure 54 Network diagram



Device	Interface	IP address	Device	Interface	IP address
Switch A	Vlan-int100	192.168.1.1/24	Switch D	Vlan-int300	192.168.3.1/24

	Vlan-int101	10.110.1.1/24		Vlan-int400	192.168.4.1/24
Switch B	Vlan-int200	192.168.2.1/24		Vlan-int103	10.110.3.2/24
	Vlan-int101	10.110.1.2/24	Source 1	-	192.168.1.100/24
	Vlan-int102	10.110.2.1/24	Source 2	-	192.168.4.100/24
Switch C	Vlan-int102	10.110.2.2/24	Receiver 1	-	192.168.2.100/24
	Vlan-int103	10.110.3.1/24	Receiver 2	-	192.168.3.100/24
	Loop0	1.1.1.1/32			

Configuration procedure

1. Configure IP addresses and unicast routing:
 - # Configure an IP address and subnet mask for each interface as per [Figure 54](#). (Details not shown.)
 - # Configure OSPF on the switches in the BIDIR-PIM domain to ensure network-layer reachability among them. (Details not shown.)

2. Enable IP multicast routing, PIM-SM, BIDIR-PIM, and IGMP:

On Switch A, enable IP multicast routing, enable PIM-SM on each interface, and enable BIDIR-PIM.

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] pim sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim sm
[SwitchA-Vlan-interface101] quit
[SwitchA] pim
[SwitchA-pim] bidir-pim enable
[SwitchA-pim] quit
```

On Switch B, enable IP multicast routing, enable PIM-SM on each interface, enable IGMP in VLAN interface 200, and enable BIDIR-PIM.

```
<SwitchB> system-view
[SwitchB] multicast routing-enable
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] igmp enable
[SwitchB-Vlan-interface200] pim sm
[SwitchB-Vlan-interface200] quit
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] pim sm
[SwitchB-Vlan-interface101] quit
[SwitchB] interface vlan-interface 102
[SwitchB-Vlan-interface102] pim sm
[SwitchB-Vlan-interface102] quit
[SwitchB] pim
[SwitchB-pim] bidir-pim enable
[SwitchB-pim] quit
```

On Switch C, enable IP multicast routing, enable PIM-SM on each interface, and enable BIDIR-PIM.

```
<SwitchC> system-view
[SwitchC] multicast routing-enable
[SwitchC] interface vlan-interface 102
[SwitchC-Vlan-interface102] pim sm
[SwitchC-Vlan-interface102] quit
[SwitchC] interface vlan-interface 103
[SwitchC-Vlan-interface103] pim sm
[SwitchC-Vlan-interface103] quit
[SwitchC] interface loopback 0
[SwitchC-LoopBack0] pim sm
[SwitchC-LoopBack0] quit
[SwitchC] pim
[SwitchC-pim] bidir-pim enable
```

On Switch D, enable IP multicast routing, enable PIM-SM on each interface, enable IGMP in VLAN interface 300, and enable BIDIR-PIM.

```
<SwitchD> system-view
[SwitchD] multicast routing-enable
[SwitchD] interface vlan-interface 300
[SwitchD-Vlan-interface300] igmp enable
[SwitchD-Vlan-interface300] pim sm
[SwitchD-Vlan-interface300] quit
[SwitchD] interface vlan-interface 400
[SwitchD-Vlan-interface400] pim sm
[SwitchD-Vlan-interface400] quit
[SwitchD] interface vlan-interface 103
[SwitchD-Vlan-interface103] pim sm
[SwitchD-Vlan-interface103] quit
[SwitchD] pim
[SwitchD-pim] bidir-pim enable
[SwitchD-pim] quit
```

3. Configure C-BSR and C-RP:

On Switch C, configure VLAN interface 102 as a C-BSR, and loopback interface 0 as a C-RP for the entire BIDIR-PIM domain.

```
[SwitchC-pim] c-bsr vlan-interface 102
[SwitchC-pim] c-rp loopback 0 bidir
[SwitchC-pim] quit
```

4. Verify the configuration:

Display the DF information of BIDIR-PIM on Switch A.

```
[SwitchA] display pim df-info
VPN-Instance: public net
```

```
RP Address: 1.1.1.1
```

Interface	State	DF-Pref	DF-Metric	DF-Uptime	DF-Address
Vlan100	Win	100	2	01:08:50	192.168.1.1 (local)
Vlan101	Lose	100	1	01:07:49	10.110.1.2

Display the DF information of BIDIR-PIM on Switch B.

```
[SwitchB] display pim df-info
```

```
VPN-Instance: public net
```

```
RP Address: 1.1.1.1
```

Interface	State	DF-Pref	DF-Metric	DF-Uptime	DF-Address
Vlan200	Win	100	1	01:24:09	192.168.2.1 (local)
Vlan101	Win	100	1	01:24:09	10.110.1.2 (local)
Vlan102	Lose	0	0	01:23:12	10.110.2.2

Display the DF information of BIDIR-PIM on Switch C.

```
[SwitchC] display pim df-info
```

```
VPN-Instance: public net
```

```
RP Address: 1.1.1.1
```

Interface	State	DF-Pref	DF-Metric	DF-Uptime	DF-Address
Loop0	-	-	-	-	-
Vlan102	Win	0	0	01:06:07	10.110.2.2 (local)
Vlan103	Win	0	0	01:06:07	10.110.3.1 (local)

Display the DF information of BIDIR-PIM on Switch D.

```
[SwitchD] display pim df-info
```

```
VPN-Instance: public net
```

```
RP Address: 1.1.1.1
```

Interface	State	DF-Pref	DF-Metric	DF-Uptime	DF-Address
Vlan300	Win	100	1	01:19:53	192.168.3.1 (local)
Vlan400	Win	100	1	00:39:34	192.168.4.1 (local)
Vlan103	Lose	0	0	01:21:40	10.110.3.1

Display the DF information of the multicast forwarding table on Switch A.

```
[SwitchA] display multicast forwarding-table df-info
```

```
Multicast DF information of VPN-Instance: public net
```

```
Total 1 RP
```

```
Total 1 RP matched
```

```
00001. RP Address: 1.1.1.1
```

```
  MID: 0, Flags: 0x2100000:0
```

```
  Uptime: 00:08:32
```

```
  RPF interface: Vlan-interface101
```

```
  List of 1 DF interfaces:
```

```
    1: Vlan-interface100
```

Display the DF information of the multicast forwarding table on Switch B.

```
[SwitchB] display multicast forwarding-table df-info
```

```
Multicast DF information of VPN-Instance: public net
```

```
Total 1 RP
```

```
Total 1 RP matched
```

```

00001. RP Address: 1.1.1.1
      MID: 0, Flags: 0x2100000:0
      Uptime: 00:06:24
      RPF interface: Vlan-interface102
      List of 2 DF interfaces:
        1: Vlan-interface101
        2: Vlan-interface200

# Display the DF information of the multicast forwarding table on Switch C.
[SwitchC] display multicast forwarding-table df-info
Multicast DF information of VPN-Instance: public net
Total 1 RP

Total 1 RP matched

00001. RP Address: 1.1.1.1
      MID: 0, Flags: 0x2100000:0
      Uptime: 00:07:21
      RPF interface: LoopBack0
      List of 2 DF interfaces:
        1: Vlan-interface102
        2: Vlan-interface103

# Display the DF information of the multicast forwarding table on Switch D.
[SwitchD] display multicast forwarding-table df-info
Multicast DF information of VPN-Instance: public net
Total 1 RP

Total 1 RP matched

00001. RP Address: 1.1.1.1
      MID: 0, Flags: 0x2100000:0
      Uptime: 00:05:12
      RPF interface: Vlan-interface103
      List of 2 DF interfaces:
        1: Vlan-interface300
        2: Vlan-interface400

```

PIM-SSM configuration example

Network requirements

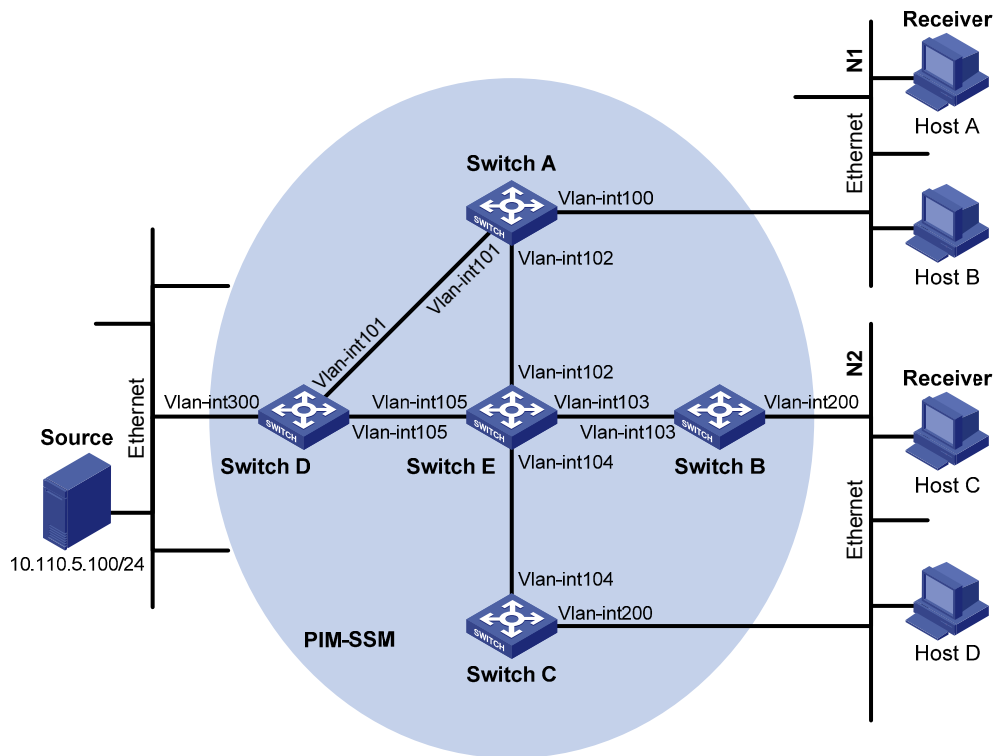
As shown in [Figure 55](#), receivers receive VOD information through multicast. The receiver groups of different organizations form stub networks, and one or more receiver hosts exist in each stub network. The entire PIM domain is operating in the SSM mode.

Host A and Host C are multicast receivers in two stub networks.

The SSM group range is 232.1.1.0/24.

IGMPv3 runs between Switch A and N1 and between Switch B/Switch C and N2.

Figure 55 Network diagram



Device	Interface	IP address	Device	Interface	IP address
Switch A	Vlan-int100	10.110.1.1/24	Switch D	Vlan-int300	10.110.5.1/24
	Vlan-int101	192.168.1.1/24		Vlan-int101	192.168.1.2/24
	Vlan-int102	192.168.9.1/24		Vlan-int105	192.168.4.2/24
Switch B	Vlan-int200	10.110.2.1/24	Switch E	Vlan-int104	192.168.3.2/24
	Vlan-int103	192.168.2.1/24		Vlan-int103	192.168.2.2/24
Switch C	Vlan-int200	10.110.2.2/24		Vlan-int102	192.168.9.2/24
	Vlan-int104	192.168.3.1/24		Vlan-int105	192.168.4.1/24

Configuration procedure

1. Configure IP addresses and unicast routing:
Configure the IP address and subnet mask for each interface as per Figure 55. (Details not shown.)
Configure OSPF on the switches in the PIM-SSM domain to ensure network-layer reachability among them. (Details not shown.)
2. Enable IP multicast routing, and enable PIM-SM and IGMP:
Enable IP multicast routing on Switch A, enable PIM-SM on each interface, and run IGMPv3 on VLAN-interface 100, which connects Switch A to the stub network.

```

<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] igmp enable
[SwitchA-Vlan-interface100] igmp version 3
[SwitchA-Vlan-interface100] pim sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101

```



```
[SwitchA-Vlan-interface101] pim sm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 102
[SwitchA-Vlan-interface102] pim sm
[SwitchA-Vlan-interface102] quit
```

The configuration on Switch B and Switch C is similar to that on Switch A. The configuration on Switch D and Switch E is also similar to that on Switch A except that it is not necessary to enable IGMP on the corresponding interfaces on these two switches.

3. Configure the SSM group range:

Configure the SSM group range to be 232.1.1.0/24 on Switch A.

```
[SwitchA] acl number 2000
[SwitchA-acl-basic-2000] rule permit source 232.1.1.0 0.0.0.255
[SwitchA-acl-basic-2000] quit
[SwitchA] pim
[SwitchA-pim] ssm-policy 2000
[SwitchA-pim] quit
```

The configuration on Switch B, Switch C, Switch D and Switch E is similar to that on Switch A.

4. Verify the configuration:

Display PIM configuration information on Switch A.

```
[SwitchA] display pim interface
VPN-Instance: public net
Interface           NbrCnt  HelloInt   DR-Pri     DR-Address
Vlan100             0        30         1          10.110.1.1    (local)
Vlan101             1        30         1          192.168.1.2
Vlan102             1        30         1          192.168.9.2
```

Assume that Host A needs to receive the information a specific multicast source S (10.110.5.100/24) sends to multicast group G (232.1.1.1). Switch A builds an SPT toward the multicast source. Switches on the SPT path (Switch A and Switch D) have generated an (S, G) entry, but Switch E, which is not on the SPT path, does not have multicast routing entries. You can use the **display pim routing-table** command to view the PIM routing table information on each switch. For example:

Display PIM routing table information on Switch A.

```
[SwitchA] display pim routing-table
VPN-Instance: public net
Total 0 (*, G) entry; 1 (S, G) entry

(10.110.5.100, 232.1.1.1)
  Protocol: pim-ssm, Flag:
  UpTime: 00:13:25
  Upstream interface: Vlan-interface101
    Upstream neighbor: 192.168.1.2
    RPF prime neighbor: 192.168.1.2
  Downstream interface(s) information:
  Total number of downstreams: 1
    1: Vlan-interface100
      Protocol: igmp, UpTime: 00:13:25, Expires: 00:03:25
```

Display PIM routing table information on Switch D.

```

[SwitchD] display pim routing-table
VPN-Instance: public net
Total 0 (*, G) entry; 1 (S, G) entry

(10.110.5.100, 232.1.1.1)
  Protocol: pim-ssm, Flag: LOC
  UpTime: 00:12:05
  Upstream interface: Vlan-interface300
    Upstream neighbor: NULL
    RPF prime neighbor: NULL
  Downstream interface(s) information:
  Total number of downstreams: 1
    1: Vlan-interface105
      Protocol: pim-ssm, UpTime: 00:12:05, Expires: 00:03:25

```

Troubleshooting PIM

A multicast distribution tree cannot be built correctly

Symptom

None of the routers in the network (including routers directly connected with multicast sources and receivers) have multicast forwarding entries. That is, a multicast distribution tree cannot be built correctly and clients cannot receive multicast data.

Analysis

- When PIM-DM runs on the entire network, multicast data is flooded from the first hop router connected with the multicast source to the last hop router connected with the clients. When the multicast data is flooded to a router, regardless of which router it is, the router creates (S, G) entries only if it has a route to the multicast source. If the router does not have a route to the multicast source, or if PIM-DM is not enabled on the router's RPF interface to the multicast source, the router cannot create (S, G) entries.
- When PIM-SM runs on the entire network and when a router will join the SPT, the router creates (S, G) entries only if it has a route to the multicast source. If the router does not have a route to the multicast source, or if PIM-DM is not enabled on the router's RPF interface to the multicast source, the router cannot create (S, G) entries.
- When a multicast router receives a multicast packet, it searches the existing unicast routing table for the optimal route to the RPF check object. The outgoing interface of this route will act as the RPF interface and the next hop will be taken as the RPF neighbor. The RPF interface completely relies on the existing unicast route, and is independent of PIM. The RPF interface must be PIM-enabled, and the RPF neighbor must also be a PIM neighbor. If PIM is not enabled on the router where the RPF interface or the RPF neighbor resides, the establishment of a multicast distribution tree will surely fail, causing abnormal multicast forwarding.
- Because a hello message does not carry the PIM mode information, a router that is running PIM cannot identify what PIM mode its PIM neighbor is running. If different PIM modes are enabled on the RPF interface and on the corresponding interface of the RPF neighbor router, the establishment of a multicast distribution tree will fail, causing abnormal multicast forwarding.
- The same PIM mode must run on the entire network. Otherwise, the establishment of a multicast distribution tree will fail, causing abnormal multicast forwarding.

Solution

1. Use the **display ip routing-table** command to verify that a unicast route exists from the receiver host to the multicast source.
2. Use the **display pim interface** command to verify that PIM is enabled on the interfaces, especially on the RPF interface. If PIM is not enabled on the interface, use the **pim dm** or **pim sm** command to enable PIM-DM or PIM-SM.
3. Use the **display pim neighbor** command to verify that the RPF neighbor is a PIM neighbor.
4. Verify that PIM and IGMP are enabled on the interfaces directly connecting to the multicast source and to the receivers.
5. Use the **display pim interface verbose** command to verify that the same PIM mode is enabled on the RPF interface and the corresponding interface of the RPF neighbor router.
6. Verify that the same PIM mode is enabled on all the routers in the entire network. Make sure that the same PIM mode is enabled on all the routers: PIM-SM on all routers, or PIM-DM on all routers. In the case of PIM-SM, also check that the BSR and RP configurations are correct.

Multicast data abnormally terminated on an intermediate router

Symptom

An intermediate router can receive multicast data successfully, but the data cannot reach the last hop router. An interface on the intermediate router receives data but no corresponding (S, G) entry is created in the PIM routing table.

Analysis

- If a multicast forwarding boundary has been configured through the **multicast boundary** command, any multicast packet will be kept from crossing the boundary, and no routing entry can be created in the PIM routing table.
- In addition, the **source-policy** command filters received multicast packets. If the multicast data fails to pass the ACL rule defined in this command, PIM cannot create the route entry either.

Solution

1. Use the **display current-configuration** command to verify the multicast forwarding boundary settings. Use the **multicast boundary** command to change the multicast forwarding boundary settings.
2. Use the **display current-configuration** command to verify the multicast filter configuration. Change the ACL rule defined in the **source-policy** command so that the source/group address of the multicast data can pass ACL filtering.

RPs cannot join SPT in PIM-SM

Symptom

An RPT cannot be established correctly, or the RPs cannot join the SPT to the multicast source.

Analysis

- As the core of a PIM-SM domain, the RPs serve specific multicast groups. Multiple RPs can coexist in a network. Make sure that the RP information on all routers is exactly the same and that a specific group is mapped to the same RP. Otherwise, multicast forwarding will fail.
- If the static RP mechanism is used, the same static RP command must be executed on all the routers in the entire network. Otherwise, multicast forwarding will fail.

Solution

1. Use the **display ip routing-table** command to verify that a route is available on each router to the RP.
2. Use the **display pim rp-info** command to verify that the RP information is consistent on all routers.
3. Use the **display pim rp-info** command to verify that the same static RP address has been configured on all the routers in the entire network.

RPT establishment failure or source registration failure in PIM-SM

Symptom

C-RPs cannot unicast advertise messages to the BSR. The BSR does not advertise bootstrap messages containing C-RP information and has no unicast route to any C-RP. An RPT cannot be established correctly, or the DR cannot perform source registration with the RP.

Analysis

- The C-RPs periodically send C-RP-Adv messages to the BSR by unicast. If a C-RP has no unicast route to the BSR, the BSR cannot receive C-RP-Adv messages from that C-RP and the bootstrap message of the BSR will not contain the information of that C-RP.
- In addition, if the BSR does not have a unicast route to a C-RP, it will discard the C-RP-Adv messages from that C-RP, and therefore the bootstrap messages of the BSR will not contain the information of that C-RP.
- The RP is the core of a PIM-SM domain. Make sure that the RP information on all routers is exactly the same, a specific group G is mapped to the same RP, and unicast routes are available to the RP.

Solution

1. Use the **display ip routing-table** command to verify that routes are available on each router to the RP and the BSR and whether a route is available between the RP and the BSR. Make sure that each C-RP has a unicast route to the BSR, the BSR has a unicast route to each C-RP, and all the routers in the entire network have a unicast route to the RP.
2. PIM-SM needs the support of the RP and BSR. Use the **display pim bsr-info** command to verify that the BSR information is available on each router, and then use the **display pim rp-info** command to verify that the RP information is correct.
3. Use the **display pim neighbor** command to verify that the normal PIM neighboring relationships have been established among the routers.

Configuring MSDP (available only on the HP 5500 EI)

Overview

Multicast source discovery protocol (MSDP) is an inter-domain multicast solution that addresses the interconnection of protocol independent multicast sparse mode (PIM-SM) domains. You can use it to discover multicast source information in other PIM-SM domains.

In the basic PIM-SM mode, a multicast source registers only with the RP in the local PIM-SM domain, and the multicast source information about a domain is isolated from that of another domain. As a result, the RP obtains the source information only within the local domain, and a multicast distribution tree is built only within the local domain to deliver multicast data from a local multicast source to local receivers. MSDP enables the RPs of different PIM-SM domains to share their multicast source information, so that the local RP can join multicast sources in other domains, and multicast data can be transmitted among different domains.

With MSDP peer relationships established between appropriate routers in the network, the RPs of different PIM-SM domains are interconnected with one another. These MSDP peers exchange source active (SA) messages, so that the multicast source information is shared among these different domains.

NOTE:

- MSDP is applicable only if the intra-domain multicast protocol is PIM-SM.
 - MSDP is meaningful only for the any-source multicast (ASM) model.
-

For more information about the concepts of designated router (DR), bootstrap router (BSR), candidate-BSR (C-BSR), rendezvous point (RP), candidate-RP (C-RP), shortest path tree (SPT) and rendezvous point tree (RPT) mentioned in this document, see "[Configuring PIM \(available only on the HP 5500 EI\)](#)"

The term "router" in this document refers to both routers and Layer 3 switches.

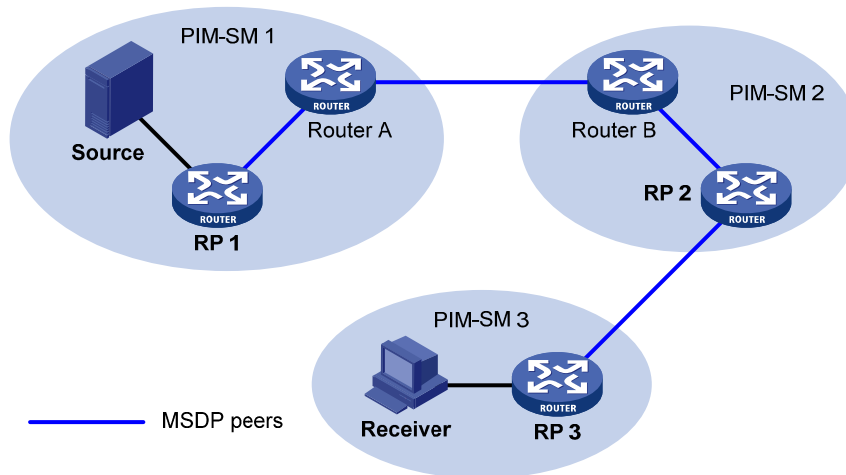
The term "interface" in the MSDP features refers to Layer 3 interfaces, including VLAN interfaces and route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).

How MSDP works

MSDP peers

Configuring one or more pairs of MSDP peers in the network forms an MSDP interconnection map, where the RPs of different PIM-SM domains are interconnected in series. An SA message that an RP sends and that these MSDP peers relay can be delivered to all other RPs.

Figure 56 Where MSDP peers are in the network



As shown in Figure 56, an MSDP peer can be created on any PIM-SM router. MSDP peers created on PIM-SM routers that assume different roles function differently.

1. MSDP peers on RPs include the following types:
 - **Source-side MSDP peer**—The MSDP peer nearest to the multicast source (Source), typically the source-side RP, like RP 1. The source-side RP creates SA messages and sends the messages to its remote MSDP peer to notify the MSDP peer of the locally registered multicast source information. A source-side MSDP peer must be created on the source-side RP. Otherwise it will not be able to advertise the multicast source information out of the PIM-SM domain.
 - **Receiver-side MSDP peer**—The MSDP peer nearest to the receivers, typically the receiver-side RP, like RP 3. After receiving an SA message, the receiver-side MSDP peer resolves the multicast source information carried in the message and joins the SPT rooted at the source across the PIM-SM domain. When multicast data from the multicast source arrives, the receiver-side MSDP peer forwards the data to the receivers along the RPT.
 - **Intermediate MSDP peer**—An MSDP peer with multicast remote MSDP peers, like RP 2. An intermediate MSDP peer forwards SA messages received from one remote MSDP peer to other remote MSDP peers, functioning as a relay of multicast source information.
2. MSDP peers created on common PIM-SM routers (other than RPs)

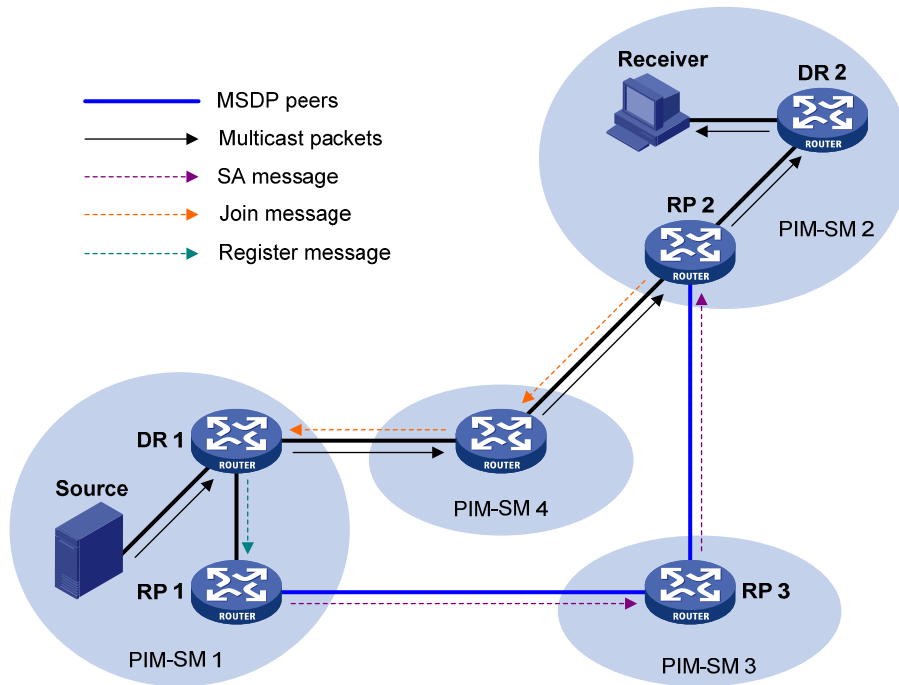
Router A and Router B are MSDP peers on common multicast routers. Such MSDP peers just forward received SA messages.

In a PIM-SM network running the BSR mechanism, the RP is dynamically elected from C-RPs. To enhance network robustness, a PIM-SM network typically has more than one C-RP. As the RP election result is unpredictable, MSDP peering relationships must be built among all C-RPs so that the winner C-RP is always on the "MSDP interconnection map," and loser C-RPs will assume the role of common PIM-SM routers on the "MSDP interconnection map."

Inter-domain multicast delivery through MSDP

As shown in Figure 57, an active source (Source) exists in the domain PIM-SM 1, and RP 1 has learned the existence of Source through multicast source registration. If RPs in PIM-SM 2 and PIM-SM 3 also seek the specific location of Source so that receiver hosts can receive multicast traffic that the source sends, HP recommends you to establish MSDP peering relationships between RP 1 and RP 3 and between RP 3 and RP 2, respectively.

Figure 57 Inter-domain multicast delivery through MSDP



The process of implementing PIM-SM inter-domain multicast delivery by leveraging MSDP peers is as follows:

1. When the multicast source in PIM-SM 1 sends the first multicast packet to multicast group G, DR 1 encapsulates the multicast data within a register message and sends the register message to RP 1. Then, RP 1 identifies the information related to the multicast source.
2. As the source-side RP, RP 1 creates SA messages and periodically sends the SA messages to its MSDP peer. An SA message contains the source address (S), the multicast group address (G), and the address of the RP that has created this SA message (namely, RP 1).
3. On MSDP peers, each SA message undergoes a reverse path forwarding (RPF) check and multicast policy-based filtering, so that only SA messages that have arrived along the correct path and passed the filtering are received and forwarded. This avoids delivery loops of SA messages. In addition, you can configure MSDP peers into an MSDP mesh group so as to avoid flooding of SA messages between MSDP peers.

An MSDP mesh group refers to a group of MSDP peers that have MSDP peering relationships among one another and share the same group name.

4. SA messages are forwarded from one MSDP peer to another, and finally the information about the multicast source traverses all PIM-SM domains with MSDP peers (PIM-SM 2 and PIM-SM 3, in this example).
5. After receiving the SA message that RP 1 created, RP 2 in PIM-SM 2 determines whether any receivers for the multicast group exist in the domain.
 - If receivers for the multicast group exist in the domain, the RPT for the multicast group G is maintained between RP 2 and the receivers. RP 2 creates an (S, G) entry and sends an (S, G) join message hop by hop toward DR 1 at the multicast source side, so that it can directly join the SPT rooted at the source over other PIM-SM domains. Then, the multicast data can flow along the SPT to RP 2 and RP 2 can forward the data to the receivers along the RPT. After receiving the multicast traffic, the DR at the receiver side (DR 2) determines whether to initiate an RPT-to-SPT switchover process.

- If no receivers for the group exist in the domain, RP 2 neither creates an (S, G) entry nor joins the SPT rooted at the source.

NOTE:

When using MSDP for inter-domain multicasting, once an RP receives information from a multicast source, it no longer relies on RPs in other PIM-SM domains. The receivers can override the RPs in other domains and directly join the multicast source-based SPT.

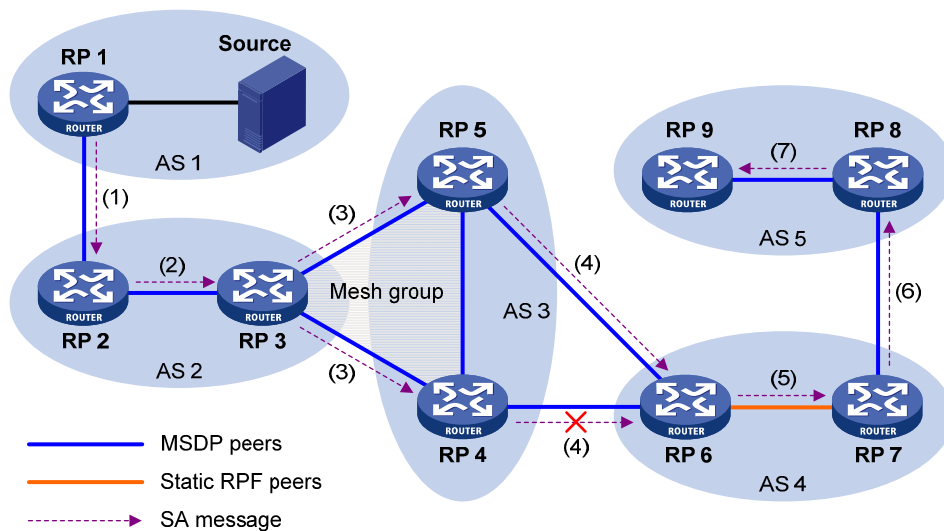
RPF check rules for SA messages

As shown in Figure 58, the autonomous systems in the network are AS 1 through AS 5, with IGP enabled on routers within each AS and BGP or MBGP as the interoperation protocol among different ASs. Each AS contains at least one PIM-SM domain, and each PIM-SM domain contains one or more RPs. MSDP peering relationships have been established among different RPs. RP 3, RP 4, and RP 5 are in an MSDP mesh group. On RP 7, RP 6 is configured as its static RPF peer.

NOTE:

When an RP receives an SA message from a static RPF peer, the RP accepts the SA message and forwards it to other peers without performing an RPF check.

Figure 58 Diagram for RPF check for SA messages



As shown in Figure 58, these MSDP peers dispose of SA messages according to the following RPF check rules:

1. When RP 2 receives an SA message from RP 1:

Because the source-side RP address carried in the SA message is the same as the MSDP peer address, which means that the MSDP peer where the SA is from is the RP that has created the SA message, RP 2 accepts the SA message and forwards it to its other MSDP peer (RP 3).
2. When RP 3 receives the SA message from RP 2:

Because the SA message is from an MSDP peer (RP 2) in the same AS, and the MSDP peer is the next hop on the optimal path to the source-side RP, RP 3 accepts the message and forwards it to other peers (RP 4 and RP 5).
3. When RP 4 and RP 5 receive the SA message from RP 3:

Because the SA message is from an MSDP peer (RP 3) in the same mesh group, RP 4 and RP 5 both accept the SA message, but they do not forward the message to other members in the mesh group. Instead, they forward it to other MSDP peers (RP 6 in this example) out of the mesh group.

4. When RP 6 receives the SA messages from RP 4 and RP 5 (suppose RP 5 has a higher IP address):
Although RP 4 and RP 5 are in the same AS (AS 3) and both are MSDP peers of RP 6, because RP 5 has a higher IP address, RP 6 accepts only the SA message from RP 5.
5. When RP 7 receives the SA message from RP 6:
Because the SA message is from a static RPF peer (RP 6), RP 7 accepts the SA message and forwards it to other peer (RP 8).
6. When RP 8 receives the SA message from RP 7:
A BGP or MBGP route exists between two MSDP peers in different ASs. Because the SA message is from an MSDP peer (RP 7) in a different AS, and the MSDP peer is the next hop on the BGP or MBGP route to the source-side RP, RP 8 accepts the message and forwards it to its other peer (RP 9).
7. When RP 9 receives the SA message from RP 8:
Because RP 9 has only one MSDP peer, RP 9 accepts the SA message.

SA messages from paths other than those described previously are not accepted or forwarded by MSDP peers.

Intra-domain Anycast RP through MSDP

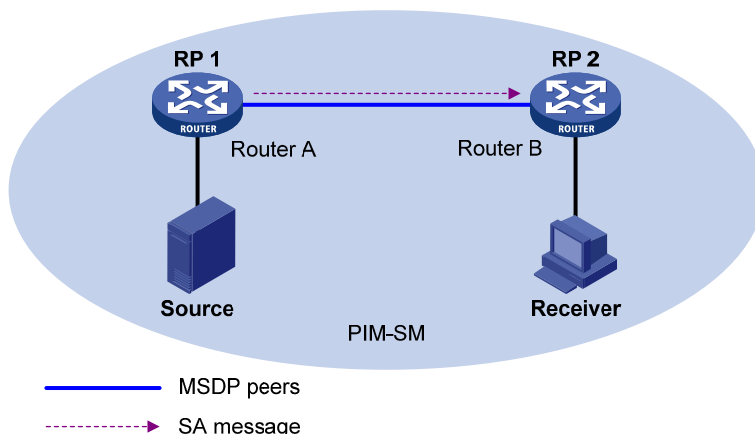
Anycast RP refers to an application that enables load balancing and redundancy backup between two or more RPs within a PIM-SM domain by configuring the same IP address for, and establishing MSDP peering relationships between, these RPs.

Usually an Anycast RP address is configured on a logic interface, like a loopback interface. An MSDP peer address must be different from the Anycast RP address.

Be sure to configure a 32-bit subnet mask (255.255.255.255) for the Anycast RP address sure, which means configure the Anycast RP address into a host address.

As shown in Figure 59, within a PIM-SM domain, a multicast source sends multicast data to multicast group G, and Receiver is a member of the multicast group. To implement Anycast RP, configure the same IP address (known as Anycast RP address, typically a private address) on Router A and Router B, configure these interfaces as C-RPs, and establish an MSDP peering relationship between Router A and Router B.

Figure 59 Intra-domain Anycast RP through MSDP



The work process of Anycast RP is as follows:

1. The multicast source registers with the nearest RP. In this example, Source registers with RP 1, with its multicast data encapsulated in the register message. When the register message arrives at RP 1, RP 1 de-encapsulates the message.
2. Receivers send join messages to the nearest RP to join in the RPT rooted at this RP. In this example, Receiver joins the RPT rooted at RP 2.
3. RPs share the registered multicast information by means of SA messages. In this example, RP 1 creates an SA message and sends it to RP 2, with the multicast data from Source encapsulated in the SA message. When the SA message reaches RP 2, RP 2 de-encapsulates the message.
4. Receivers receive the multicast data along the RPT and directly join the SPT rooted at the multicast source. In this example, RP 2 forwards the multicast data down the RPT. When Receiver receives the multicast data from Source, it directly joins the SPT rooted at Source.

The significance of Anycast RP is as follows:

- **Optimal RP path**—A multicast source registers with the nearest RP so that an SPT with the optimal path is built. A receiver joins the nearest RP so that an RPT with the optimal path is built.
- **Load balancing between RPs**—Each RP maintains just part of the source/group information within the PIM-SM domain and forward part of the multicast data, thereby achieving load balancing between different RPs.
- **Redundancy backup between RPs**—When an RP fails, the multicast source that previously registered with the RP or the receivers that previously joined the RP will register with or join another nearest RP, thereby achieving redundancy backup between RPs.

MSDP support for VPNs

The interfaces on the multicast routers in a VPN can set up MSDP peering relationships between each other. By exchanging SA messages between MSDP peers, multicast transmission in a VPN between different PIM-SM domains can be implemented.

To support MSDP for VPNs, a multicast router that runs MSDP maintains an independent set of MSDP mechanism for each VPN that it supports, including SA cache, peering connection, timers, sending cache, and cache for exchanging PIM messages. The information in one VPN is isolated from another, and MSDP and PIM-SM messages can be exchanged only within the same VPN.

Protocols and standards

- RFC 3618, *Multicast Source Discovery Protocol (MSDP)*
- RFC 3446, *Anycast Rendezvous Point (RP) mechanism using Protocol Independent Multicast (PIM) and Multicast Source Discovery Protocol (MSDP)*

MSDP configuration task list

Task	Remarks
Configuring basic MSDP functions	Enabling MSDP Required
	Creating an MSDP peer connection Required
	Configuring a static RPF peer Optional
Configuring an MSDP peer	Configuring MSDP peer description Optional

Task		Remarks
connection	Configuring an MSDP mesh group	Optional
	Configuring MSDP peer connection control	Optional
	Configuring SA message content	Optional
Configuring SA messages related parameters	Configuring SA request messages	Optional
	Configuring SA message filtering rules	Optional
	Configuring the SA cache mechanism	Optional

Configuring basic MSDP functions

ⓘ IMPORTANT:

All the configuration tasks should be carried out on RPs in PIM-SM domains, and each of these RPs acts as an MSDP peer.

Configuration prerequisites

Before you configure basic MSDP functions, complete the following tasks:

- Configure any unicast routing protocol so that all devices in the domain are interoperable at the network layer.
- Configure PIM-SM to enable intra-domain multicast forwarding.
- Determine the IP addresses of MSDP peers.
- Determine the address prefix list for an RP address filtering policy.

Enabling MSDP

Enabling MSDP globally for the public network

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IP multicast routing.	multicast routing-enable	Disabled by default.
3. Enable MSDP and enter public network MSDP view.	msdp	Disabled by default.

Enabling MSDP in a VPN instance

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a VPN instance and enter VPN instance view.	ip vpn-instance <i>vpn-instance-name</i>	N/A

Step	Command	Remarks
3.	Configure a route-distinguisher (RD) for the VPN instance. route-distinguisher <i>route-distinguisher</i>	No RD is configured by default.
4.	Enable IP multicast routing. multicast routing-enable	Disabled by default.
5.	Return to system view. quit	N/A
6.	Enable MSDP and enter VPN instance MSDP view. msdp vpn-instance <i>vpn-instance-name</i>	Disabled by default.

For more information about the **ip vpn-instance** and **route-distinguisher** commands, see *IP Routing Command Reference*.

For more information about the **multicast routing-enable** command, see *IP Multicast Command Reference*.

Creating an MSDP peer connection

An MSDP peering relationship is identified by an address pair, namely, the address of the local MSDP peer and that of the remote MSDP peer. An MSDP peer connection must be created on both devices that are a pair of MSDP peers.

To create an MSDP peer connection:

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Enter public network MSDP view or VPN instance MSDP view. msdp [vpn-instance <i>vpn-instance-name]</i>	N/A
3.	Create an MSDP peer connection. peer peer-address connect-interface <i>interface-type interface-number</i>	No MSDP peer connection is created by default.

NOTE:

If an interface of the router is shared by an MSDP peer and a BGP or MBGP peer at the same time, HP recommends you to configure the IP address of the MSDP peer the same as that of the BGP or MBGP peer.

Configuring a static RPF peer

Configuring static RPF peers avoids RPF check of SA messages.

To configure a static RPF peer:

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Enter public network MSDP view or VPN instance MSDP view. msdp [vpn-instance <i>vpn-instance-name]</i>	N/A
3.	Configure a static RPF peer. static-rpf-peer peer-address <i>[rp-policy ip-prefix-name]</i>	No static RPF peer is configured by default.

NOTE:

If only one MSDP peer is configured on a router, this MSDP will be registered as a static RPF peer.

Configuring an MSDP peer connection

Configuration prerequisites

Before you configure an MSDP peer connection, complete the following tasks:

- Configure any unicast routing protocol so that all devices in the domain are interoperable at the network layer.
- Configure basic MSDP functions.
- Determine the description of MSDP peers.
- Determine the name of an MSDP mesh group.
- Determine the MSDP peer connection retry interval.
- Determine the MD5 authentication password for the TCP connection to be established with an MSDP peer.

Configuring MSDP peer description

With the MSDP peer description information, the administrator can easily distinguish different MSDP peers to better manage MSDP peers.

To configure description for an MSDP peer:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network MSDP view or VPN instance MSDP view.	msdp [vpn-instance vpn-instance-name]	N/A
3. Configure description for an MSDP peer.	peer peer-address description text	No description is configured for an MSDP peer by default.

Configuring an MSDP mesh group

An AS can contain multiple MSDP peers. You can use the MSDP mesh group mechanism to avoid SA message flooding among these MSDP peers and optimize the multicast traffic.

An MSDP peer in an MSDP mesh group forwards SA messages (that have passed the RPF check) from outside the mesh group to the other members in the mesh group. A mesh group member accepts SA messages from inside the group without performing an RPF check, and does not forward the message within the mesh group. This mechanism not only avoids SA flooding but also simplifies the RPF check mechanism because you do not need to run BGP or MBGP between these MSDP peers.

By configuring the same mesh group name for multiple MSDP peers, you can create a mesh group that contains these MSDP peers.

! **IMPORTANT:**

- Before grouping multiple routers into an MSDP mesh group, make sure that these routers are interconnected with one another.
 - If you configure more than one mesh group name on an MSDP peer, only the last configuration is effective.
-

To create an MSDP mesh group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network MSDP view or VPN instance MSDP view.	msdp [vpn-instance vpn-instance-name]	N/A
3. Create an MSDP mesh group and assign an MSDP peer to that mesh group.	peer peer-address mesh-group name	An MSDP peer does not belong to any mesh group by default.

Configuring MSDP peer connection control

MSDP peers are interconnected over TCP (port number 639). You can flexibly control sessions between MSDP peers by manually deactivating and reactivating the MSDP peering connections. When the connection between two MSDP peers is deactivated, SA messages will no longer be delivered between them, and the TCP connection is closed without any connection setup retry. The configuration information, however, remain unchanged.

A TCP connection is required in the following situations:

- When a new MSDP peer is created
- When you reactivate a previously deactivated MSDP peer connection
- When a previously failed MSDP peer attempts to resume operation

You can adjust the interval between MSDP peering connection retries.

To enhance MSDP security, you can configure an MD5 authentication password for the TCP connection to be established with an MSDP peer. If the MD5 authentication fails, the TCP connection cannot be established.

! **IMPORTANT:**

The MSDP peers involved in the MD5 authentication must have the same authentication method and password. Otherwise, the authentication fails and the TCP connection cannot be established.

To configure MSDP peer connection control:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network MSDP view or VPN instance MSDP view.	msdp [vpn-instance vpn-instance-name]	N/A

Step	Command	Remarks
3. Deactivate an MSDP peer.	shutdown <i>peer-address</i>	Optional. Active by default.
4. Configure the interval between MSDP peer connection retries.	timer retry <i>interval</i>	Optional. 30 seconds by default.
5. Configure an MD5 authentication key for the TCP connection to be established with an MSDP peer.	peer <i>peer-address</i> password { cipher simple } <i>password</i>	Optional. By default, MD5 authentication is not performed before a TCP connection is established.

Configuring SA messages related parameters

Configuration prerequisites

Before you configure SA message delivery, complete the following tasks:

- Configure any unicast routing protocol so that all devices in the domain are interoperable at the network layer.
- Configure basic MSDP functions.
- Determine the ACL rules for filtering SA request messages.
- Determine the ACL rules as SA message creation rules.
- Determine the ACL rules for filtering SA messages to be received and forwarded.
- Determine the TTL threshold for multicast packet encapsulation in SA messages.
- Determine the maximum number of (S, G) entries learned from the specified MSDP peer that the router can cache.

Configuring SA message content

Some multicast sources send multicast data at an interval longer than the aging time of (S, G) entries. In this case, the source-side DR must encapsulate multicast data packet by packet in register messages and send them to the source-side RP. The source-side RP transmits the (S, G) information to the remote RP through SA messages. Then the remote RP joins the source-side DR and builds an SPT. Because the (S, G) entries have timed out, remote receivers can never receive the multicast data from the multicast source.

After the source-side RP is enabled to encapsulate multicast data in SA messages, if the RP wants to send a multicast packet, it encapsulates the multicast packet in an SA message and sends it. After receiving the SA message, the remote RP de-encapsulates the SA message and delivers the multicast packet to the receivers in the local domain along the RPT.

The MSDP peers deliver SA messages to one another. After receiving an SA message, a router performs RPF check on the message. If the router finds that the remote RP address is the same as the local RP address, it discards the SA message. In the Anycast RP application, however, you must configure RPs with the same IP address on two or more routers in the same PIM-SM domain and configure these routers as MSDP peers to one another. Therefore, a logic RP address (namely, the RP address on the logic interface) that is different from the actual RP address must be designated for SA messages so that the messages can pass the RPF check.

To configure the SA message content:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network MSDP view or VPN instance MSDP view.	msdp [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Enable encapsulation of multicast data in SA messages.	encap-data-enable	Optional. Disabled by default.
4. Configure the interface address as the RP address in SA messages.	originating-rp <i>interface-type</i> <i>interface-number</i>	Optional. PIM RP address by default.

Configuring SA request messages

By default, after receiving a new join message, a router does not send an SA request message to any MSDP peer. Instead, it waits for the next SA message from its MSDP peer. This will cause the receiver to delay obtaining multicast source information. To enable a new receiver to get the active multicast source information as early as possible, you can configure routers to send SA request messages to the designated MSDP peers after receiving a join message of a new receiver.

! IMPORTANT:

Before you can enable the device to send SA requests, be sure to disable the SA message cache mechanism.

To configure SA message transmission and filtering:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network MSDP view or VPN instance MSDP view.	msdp [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Enable the device to send SA request messages.	peer <i>peer-address</i> request-sa-enable	Optional. Disabled by default.
4. Configure a filtering rule for SA request messages.	peer <i>peer-address</i> sa-request-policy [acl <i>acl-number</i>]	Optional. SA request messages are not filtered by default.

Configuring SA message filtering rules

Configuration guidelines

By configuring an SA message creation rule, you can enable the router to filter the (S, G) entries to be advertised when creating an SA message, so that the propagation of messages of multicast sources is controlled.

By configuring a filtering rule for receiving or forwarding SA messages, you can enable the router to filter the (S, G) forwarding entries to be advertised when receiving or forwarding an SA message, so that the propagation of multicast source information is controlled at SA message reception or forwarding.

By configuring a TTL threshold for multicast data packet encapsulation in SA messages, you can control the multicast data packet encapsulation in SA messages and limit the propagation range of SA messages:

- Before creating an SA message with an encapsulated multicast data packet, the router checks the TTL value of the multicast data packet. If the TTL value is less than the threshold, the router does not create an SA message. If the TTL value is greater than or equal to the threshold, the router encapsulates the multicast data in an SA message and sends the SA message.
- After receiving an SA message with an encapsulated multicast data packet, the router decreases the TTL value of the multicast packet by 1 and then checks the TTL value. If the TTL value is less than the threshold, the router does not forward the SA message to the designated MSDP peer. If the TTL value is greater than or equal to the threshold, the router re-encapsulates the multicast data in an SA message and sends the SA message.

Configuration procedure

To configure a filtering rule for receiving or forwarding SA messages:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network MSDP view or VPN instance MSDP view.	msdp [vpn-instance <i>vpn-instance-name</i>]	N/A
3. Configure an SA message creation rule.	import-source [acl <i>acl-number</i>]	No restrictions on (S, G) entries by default.
4. Configure a filtering rule for receiving or forwarding SA messages.	peer <i>peer-address</i> sa-policy { import export } [acl <i>acl-number</i>]	No filtering rule by default.
5. Configure the TTL threshold for multicast data packet encapsulation in SA messages.	peer <i>peer-address</i> minimum-ttl <i>tll-value</i>	Optional. 0 by default.

Configuring the SA cache mechanism

To reduce the time spent in obtaining the multicast information, you can enable the SA cache mechanism to cache (S, G) entries contained in SA messages locally on the router. However, caching (S, G) entries uses memory space on the router.

When the SA cache mechanism is enabled and the router receives a new (*, G) join message, the router searches its SA cache first.

- If the corresponding (S, G) entry does not exist in the cache, the router waits for the SA message that its MSDP peer will send in the next cycle.
- If the corresponding (S, G) entry exists in the cache, the router joins the corresponding SPT rooted at S.

To protect the router effectively against denial of service (DoS) attacks, you can set a limit on the number of (S, G) entries the router can cache.

To configure the SA message cache:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter public network MSDP view or VPN instance MSDP view.	msdp [vpn-instance vpn-instance-name]	N/A
3. Enable the SA cache mechanism.	cache-sa-enable	Optional. Enabled by default.
4. Configure the maximum number of (S, G) entries learned from the specified MSDP peer that the router can cache.	peer peer-address sa-cache-maximum sa-limit	Optional. 8192 by default.

Displaying and maintaining MSDP

Step	Command	Remarks
1. Display brief information about MSDP peers.	display msdp [all-instance vpn-instance vpn-instance-name] brief [state { connect down listen shutdown up }] [{ begin exclude include } regular-expression]	Available in any view
2. Display detailed information about the status of MSDP peers.	display msdp [all-instance vpn-instance vpn-instance-name] peer-status [peer-address] [{ begin exclude include } regular-expression]	Available in any view
3. Display the (S, G) entry information in the SA cache.	display msdp [all-instance vpn-instance vpn-instance-name] sa-cache [group-address source-address as-number] * [{ begin exclude include } regular-expression]	Available in any view
4. Display the number of (S, G) entries in the SA cache.	display msdp [all-instance vpn-instance vpn-instance-name] sa-count [as-number] [{ begin exclude include } regular-expression]	Available in any view
5. Reset the TCP connection with an MSDP peer.	reset msdp [all-instance vpn-instance vpn-instance-name] peer [peer-address]	Available in user view
6. Clear (S, G) entries in the SA cache.	reset msdp [all-instance vpn-instance vpn-instance-name] sa-cache [group-address]	Available in user view
7. Clear statistics for an MSDP peer.	reset msdp [all-instance vpn-instance vpn-instance-name] statistics [peer-address]	Available in user view

MSDP configuration examples

PIM-SM Inter-domain multicast configuration

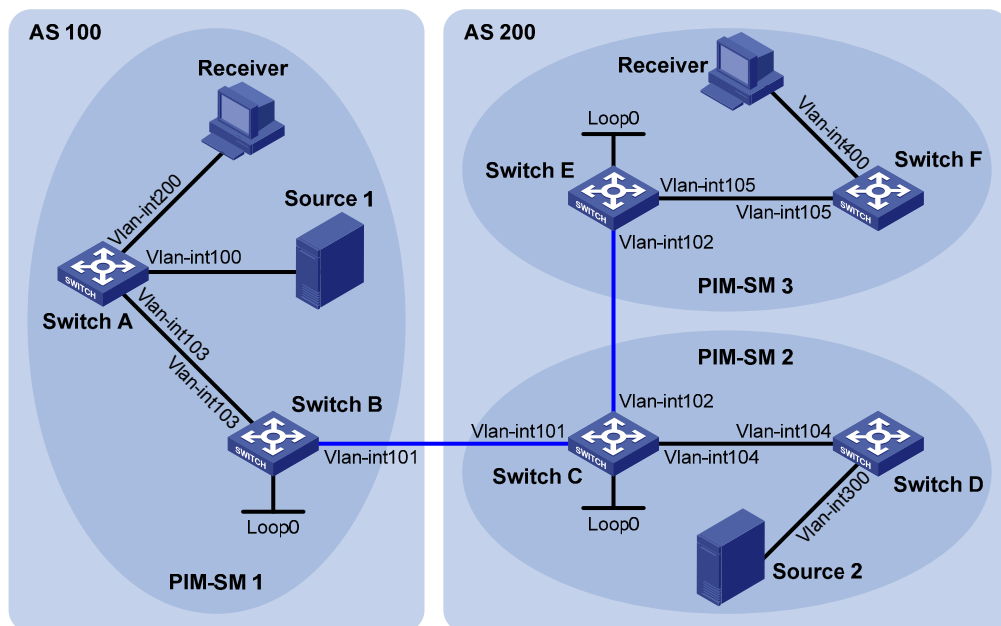
Network requirements

As shown in Figure 60, AS 100 and AS 200 run OSPF within each AS, and run BGP between each other. PIM-SM 1 belongs to AS 100, and PIM-SM 2 and PIM-SM 3 belong to AS 200. Each PIM-SM domain has at least one multicast source or receiver.

Loopback 0 is configured as the C-BSR and C-RP of the related PIM-SM domain on Switch B, Switch C, and Switch E, respectively.

An MSDP peering relationship is set up between the RPs of the PIM-SM domains to share multicast source information among the PIM-SM domains.

Figure 60 Network diagram



— MSDP peers

Device	Interface	IP address	Device	Interface	IP address
Switch A	Vlan-int103	10.110.1.2/24	Switch D	Vlan-int104	10.110.4.2/24
	Vlan-int100	10.110.2.1/24		Vlan-int300	10.110.5.1/24
	Vlan-int200	10.110.3.1/24	Switch E	Vlan-int105	10.110.6.1/24
Switch B	Vlan-int103	10.110.1.1/24		Vlan-int102	192.168.3.2/24
	Vlan-int101	192.168.1.1/24		Loop0	3.3.3.3/32
	Loop0	1.1.1.1/32	Switch F	Vlan-int105	10.110.6.2/24
Switch C	Vlan-int104	10.110.4.1/24		Vlan-int400	10.110.7.1/24
	Vlan-int102	192.168.3.1/24	Source 1	—	10.110.2.100/24
	Vlan-int101	192.168.1.2/24	Source 2	—	10.110.5.100/24
	Loop0	2.2.2.2/32			

Configuration procedure

1. Configure IP addresses and unicast routing:

Configure the IP address and subnet mask for each interface as per [Figure 60](#). (Details not shown.)

Configure OSPF for interconnection between switches in each AS. Ensure the network-layer interoperability among each AS, and ensure the dynamic update of routing information between the switches through a unicast routing protocol. (Details not shown.)

2. Enable IP multicast routing, enable PIM-SM on each interface, and configure a PIM-SM domain border:

Enable IP multicast routing on Switch A, enable PIM-SM on each interface, and enable IGMP on the host-side interface VLAN-interface 200.

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 103
[SwitchA-Vlan-interface103] pim sm
[SwitchA-Vlan-interface103] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] pim sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 200
[SwitchA-Vlan-interface200] igmp enable
[SwitchA-Vlan-interface200] pim sm
[SwitchA-Vlan-interface200] quit
```

The configuration on Switch B, Switch C, Switch D, Switch E, and Switch F is similar to the configuration on Switch A.

Configure a PIM domain border on Switch B.

```
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] pim bsr-boundary
[SwitchB-Vlan-interface101] quit
```

The configuration on Switch C and Switch E is similar to the configuration on Switch B.

3. Configure C-BSRs and C-RPs:

Configure Loopback 0 as a C-BSR and a C-RP on Switch B.

```
[SwitchB] pim
[SwitchB-pim] c-bsr loopback 0
[SwitchB-pim] c-rp loopback 0
[SwitchB-pim] quit
```

The configuration on Switch C and Switch E is similar to the configuration on Switch B.

4. Configure BGP for mutual route redistribution between BGP and OSPF:

Configure an EBGP peer, and redistribute OSPF routes on Switch B.

```
[SwitchB] bgp 100
[SwitchB-bgp] router-id 1.1.1.1
[SwitchB-bgp] peer 192.168.1.2 as-number 200
[SwitchB-bgp] import-route ospf 1
[SwitchB-bgp] quit
```

Configure an EBGP peer, and redistribute OSPF routes on Switch C.

```
[SwitchC] bgp 200
```

```
[SwitchC-bgp] router-id 2.2.2.2
[SwitchC-bgp] peer 192.168.1.1 as-number 100
[SwitchC-bgp] import-route ospf 1
[SwitchC-bgp] quit
```

Redistribute BGP routes into OSPF on Switch B.

```
[SwitchB] ospf 1
[SwitchB-ospf-1] import-route bgp
[SwitchB-ospf-1] quit
```

Redistribute BGP routes into OSPF on Switch C.

```
[SwitchB] ospf 1
[SwitchB-ospf-1] import-route bgp
[SwitchB-ospf-1] quit
```

5. Configure MSDP peers:

Configure an MSDP peer on Switch B.

```
[SwitchB] msdp
[SwitchB-msdp] peer 192.168.1.2 connect-interface vlan-interface 101
[SwitchB-msdp] quit
```

Configure an MSDP peer on Switch C.

```
[SwitchC] msdp
[SwitchC-msdp] peer 192.168.1.1 connect-interface vlan-interface 101
[SwitchC-msdp] peer 192.168.3.2 connect-interface vlan-interface 102
[SwitchC-msdp] quit
```

Configure MSDP peers on Switch E.

```
[SwitchE] msdp
[SwitchE-msdp] peer 192.168.3.1 connect-interface vlan-interface 102
[SwitchE-msdp] quit
```

6. Verify the configuration:

Use the **display bgp peer** command to view the BGP peering relationships between the switches. For example:

Display information about BGP peering relationships on Switch B.

```
[SwitchB] display bgp peer
```

```
BGP local router ID : 1.1.1.1
Local AS number : 100
Total number of peers : 1                Peers in established state : 1

Peer                AS  MsgRcvd  MsgSent  OutQ  PrefRcv  Up/Down  State
-----
192.168.1.2         200    24      21      0      6 00:13:09  Established
```

Display information about BGP peering relationships on Switch C.

```
[SwitchC] display bgp peer
```

```
BGP local router ID : 2.2.2.2
Local AS number : 200
Total number of peers : 1                Peers in established state : 1
```

```

Peer                AS  MsgRcvd  MsgSent  OutQ  PrefRcv  Up/Down  State
192.168.1.1         100      18       16    0        1 00:12:04 Established

```

To view the BGP routing table information on the switches, use the **display bgp routing-table** command. For example:

```
# Display BGP routing table information on Switch C.
```

```
[SwitchC] display bgp routing-table
```

```
Total Number of Routes: 5
```

```
BGP Local router ID is 2.2.2.2
```

```
Status codes: * - valid, ^ - VPNv4 best, > - best, d - damped,
              h - history, i - internal, s - suppressed, S - Stale
Origin : i - IGP, e - EGP, ? - incomplete
```

	Network	NextHop	MED	LocPrf	PrefVal	Path/Ogn
* >	1.1.1.1/32	192.168.1.1	0		0	100?
* >i	2.2.2.2/32	0.0.0.0	0		0	?
* >	192.168.1.0	0.0.0.0	0		0	?
* >	192.168.1.1/32	0.0.0.0	0		0	?
* >	192.168.1.2/32	0.0.0.0	0		0	?

When the multicast source in PIM-SM 1 (Source 1) and the multicast source in PIM-SM 2 (Source 2) send multicast information, receivers in PIM-SM 1 and PIM-SM 3 can receive the multicast data. You can use the **display msdp brief** command to view the brief information of MSDP peering relationships between the switches. For example:

```
# Display brief information about MSDP peering relationships on Switch B.
```

```
[SwitchB] display msdp brief
```

```
MSDP Peer Brief Information of VPN-Instance: public net
```

Configured	Up	Listen	Connect	Shutdown	Down
1	1	0	0	0	0

Peer's Address	State	Up/Down time	AS	SA Count	Reset Count
192.168.1.2	Up	00:12:27	200	13	0

```
# Display brief information about MSDP peering relationships on Switch C.
```

```
[SwitchC] display msdp brief
```

```
MSDP Peer Brief Information of VPN-Instance: public net
```

Configured	Up	Listen	Connect	Shutdown	Down
2	2	0	0	0	0

Peer's Address	State	Up/Down time	AS	SA Count	Reset Count
192.168.3.2	Up	00:15:32	200	8	0
192.168.1.1	Up	00:06:39	100	13	0

```
# Display brief information about MSDP peering relationships on Switch E.
```

```
[SwitchE] display msdp brief
```

```
MSDP Peer Brief Information of VPN-Instance: public net
```

Configured	Up	Listen	Connect	Shutdown	Down

```

1          1          0          0          0          0
Peer's Address   State   Up/Down time   AS     SA Count   Reset Count
192.168.3.1     Up     01:07:08      200    8          0

```

Display detailed MSDP peer information on Switch B.

```

[SwitchB] display msdp peer-status
MSDP Peer Information of VPN-Instance: public net
MSDP Peer 192.168.1.2, AS 200
Description:
Information about connection status:
  State: Up
  Up/down time: 00:15:47
  Resets: 0
  Connection interface: Vlan-interface101 (192.168.1.1)
  Number of sent/received messages: 16/16
  Number of discarded output messages: 0
  Elapsed time since last connection or counters clear: 00:17:51
Information about (Source, Group)-based SA filtering policy:
  Import policy: none
  Export policy: none
Information about SA-Requests:
  Policy to accept SA-Request messages: none
  Sending SA-Requests status: disable
Minimum TTL to forward SA with encapsulated data: 0
SAs learned from this peer: 0, SA-cache maximum for the peer: none
Input queue size: 0, Output queue size: 0
Counters for MSDP message:
  Count of RPF check failure: 0
  Incoming/outgoing SA messages: 0/0
  Incoming/outgoing SA requests: 0/0
  Incoming/outgoing SA responses: 0/0
  Incoming/outgoing data packets: 0/0

```

Inter-AS multicast configuration by leveraging static RPF peers

Network requirements

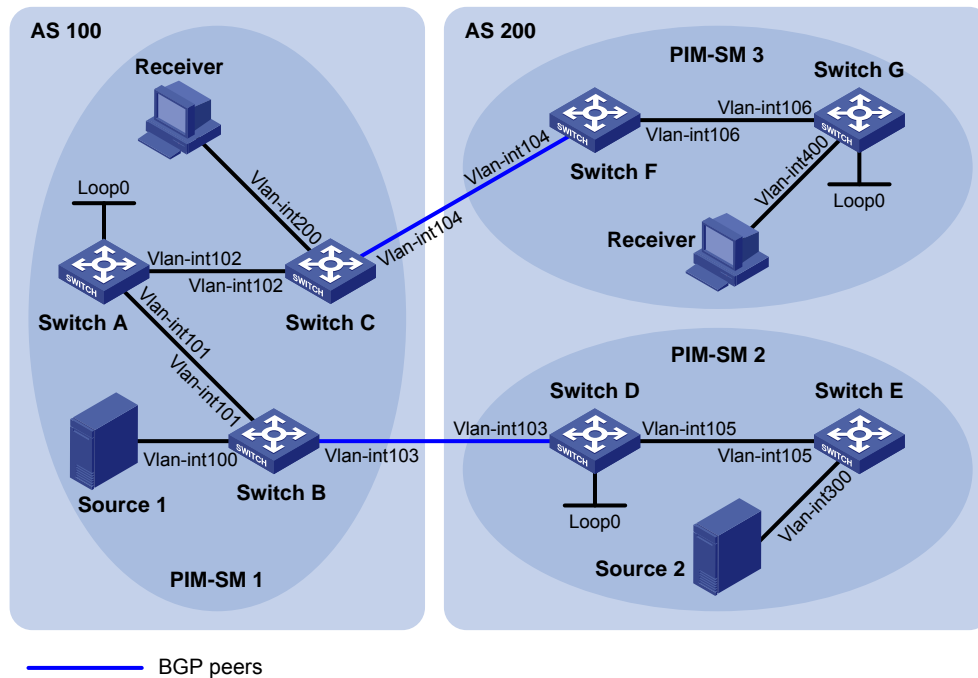
As shown in [Figure 61](#), AS 100 and AS 200 run OSPF within each AS, and run BGP between each other.

PIM-SM 1 belongs to AS 100, and PIM-SM 2 and PIM-SM 3 belong to AS 200. Each PIM-SM domain has at least one multicast source or receiver.

Loopback 0 is configured as a C-BSR and a C-RP of the related PIM-SM domain on Switch A, Switch D, and Switch G, respectively.

According to the RPF principle, the device can receive SA messages that pass the filtering policy from its static RPF peers. To share multicast source information among PIM-SM domains without changing the unicast topology structure, configure MSDP peering relationships for the RPs of the PIM-SM domains and configure static RPF peering relationships for the MSDP peers to share multicast source information among the PIM-SM domains.

Figure 61 Network diagram



Device	Interface	IP address	Device	Interface	IP address
Source 1	-	192.168.1.100/24	Switch D	Vlan-int105	10.110.5.1/24
Source 2	-	192.168.3.100/24		Vlan-int103	10.110.3.2/24
Switch A	Vlan-int101	10.110.1.1/24	Switch E	Loop0	2.2.2.2/32
	Vlan-int102	10.110.2.1/24		Vlan-int105	10.110.5.2/24
	Loop0	1.1.1.1/32		Vlan-int300	192.168.3.1/24
Switch B	Vlan-int101	10.110.1.2/24	Switch F	Vlan-int106	10.110.6.1/24
	Vlan-int100	192.168.1.1/24		Vlan-int104	10.110.4.2/24
	Vlan-int103	10.110.3.1/24		Switch G	Vlan-int106
Switch C	Vlan-int102	10.110.2.2/24	Vlan-int400		192.168.4.1/24
	Vlan-int200	192.168.2.1/24	Loop0	3.3.3.3/32	
	Vlan-int104	10.110.4.1/24			

Configuration procedure

- Configure IP addresses and unicast routing:
Configure the IP address and subnet mask for each interface as per Figure 61. (Details not shown.)
Configure OSPF for interconnection between the switches. Ensure the network-layer interoperability in each AS, and ensure the dynamic update of routing information among the switches through a unicast routing protocol. (Details not shown.)
- Enable IP multicast routing, enable PIM-SM and IGMP, and configure a PIM-SM domain border:
Enable IP multicast routing on Switch C, enable PIM-SM on each interface, and enable IGMP on the host-side interface VLAN-interface 200.

```

<SwitchC> system-view
[SwitchC] multicast routing-enable
[SwitchC] interface vlan-interface 102
[SwitchC-Vlan-interface102] pim sm
[SwitchC-Vlan-interface102] quit

```



```
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] igmp enable
[SwitchC-Vlan-interface200] pim sm
[SwitchC-Vlan-interface200] quit
[SwitchC] interface vlan-interface 104
[SwitchC-Vlan-interface104] pim sm
[SwitchC-Vlan-interface104] quit
```

The configuration on Switch A, Switch B, Switch D, Switch E, Switch F, and Switch G is similar to that on Switch C. (Details not shown.)

Configure PIM domain borders on Switch B.

```
[SwitchB] interface vlan-interface 103
[SwitchB-Vlan-interface103] pim bsr-boundary
[SwitchB-Vlan-interface103] quit
```

The configuration on Switch C, Switch D, and Switch F is similar to that on Switch B. (Details not shown.)

3. Configure C-BSRs and C-RPs:

Configure Loopback 0 as a C-BSR and a C-RP on Switch A.

```
[SwitchA] pim
[SwitchA-pim] c-bsr loopback 0
[SwitchA-pim] c-rp loopback 0
[SwitchA-pim] quit
```

The configuration on Switch D and Switch G is similar to that on Switch A. (Details not shown.)

4. Configure BGP, and redistribute BGP routing information into OSPF, and OSPF routing information into BGP:

Configure the EBGP peer, and redistribute OSPF routing information on Switch B.

```
[SwitchB] bgp 100
[SwitchB-bgp] router-id 1.1.1.2
[SwitchB-bgp] peer 10.110.3.2 as-number 200
[SwitchB-bgp] import-route ospf 1
[SwitchB-bgp] quit
```

Configure the EBGP peer, and redistribute OSPF routing information on Switch D.

```
[SwitchD] bgp 200
[SwitchD-bgp] router-id 2.2.2.2
[SwitchD-bgp] peer 10.110.3.1 as-number 100
[SwitchD-bgp] import-route ospf 1
[SwitchD-bgp] quit
```

Configure the EBGP peer, and redistribute OSPF routing information on Switch C.

```
[SwitchC] bgp 100
[SwitchC-bgp] router-id 1.1.1.3
[SwitchC-bgp] peer 10.110.4.2 as-number 200
[SwitchC-bgp] import-route ospf 1
[SwitchC-bgp] quit
```

Configure the EBGP peer, and redistribute OSPF routing information on Switch F.

```
[SwitchF] bgp 200
[SwitchF-bgp] router-id 3.3.3.1
[SwitchF-bgp] peer 10.110.4.1 as-number 100
```

```

[SwitchF-bgp] import-route ospf 1
[SwitchF-bgp] quit
# Redistribute BGP routing information into OSPF on Switch B.
[SwitchB] ospf 1
[SwitchB-ospf-1] import-route bgp
[SwitchB-ospf-1] quit
# Redistribute BGP routing information into OSPF on Switch D.
[SwitchD] ospf 1
[SwitchD-ospf-1] import-route bgp
[SwitchD-ospf-1] quit
# Redistribute BGP routing information into OSPF on Switch C.
[SwitchC] ospf 1
[SwitchC-ospf-1] import-route bgp
[SwitchC-ospf-1] quit
# Redistribute BGP routing information into OSPF on Switch F.
[SwitchF] ospf 1
[SwitchF-ospf-1] import-route bgp
[SwitchF-ospf-1] quit

```

5. Configure MSDP peers and static RPF peers:

```

# Configure Switch D and Switch G as the MSDP peers and static RPF peers of Switch A.
[SwitchA] ip ip-prefix list-dg permit 10.110.0.0 16 greater-equal 16 less-equal 32
[SwitchA] msdp
[SwitchA-msdp] peer 10.110.3.2 connect-interface vlan-interface 101
[SwitchA-msdp] peer 10.110.6.2 connect-interface vlan-interface 102
[SwitchA-msdp] static-rpf-peer 10.110.3.2 rp-policy list-dg
[SwitchA-msdp] static-rpf-peer 10.110.6.2 rp-policy list-dg
[SwitchA-msdp] quit
# Configure Switch A as the MSDP peer and static RPF peer of Switch D.
[SwitchD] ip ip-prefix list-a permit 10.110.0.0 16 greater-equal 16 less-equal 32
[SwitchD] msdp
[SwitchD-msdp] peer 10.110.1.1 connect-interface vlan-interface 103
[SwitchD-msdp] static-rpf-peer 10.110.1.1 rp-policy list-a
[SwitchD-msdp] quit
# Configure Switch A as the MSDP peer and static RPF peer of Switch G.
[SwitchG] ip ip-prefix list-a permit 10.110.0.0 16 greater-equal 16 less-equal 32
[SwitchG] msdp
[SwitchG-msdp] peer 10.110.2.1 connect-interface vlan-interface 106
[SwitchG-msdp] static-rpf-peer 10.110.2.1 rp-policy list-a
[SwitchG-msdp] quit

```

6. Verify the configuration:

Use the **display bgp peer** command to view the BGP peering relationships between the switches. If the command gives no output information on Switch A, it means that no BGP peering relationship has been established between Switch A and Switch D, or between Switch A and Switch G.

When the multicast source in PIM-SM 1 (Source 1) and the multicast source in PIM-SM 2 (Source 2) send multicast information, receivers in PIM-SM 1 and PIM-SM 3 can receive the multicast data.

You can use the **display msdp brief** command to view the brief information of MSDP peering relationships between the switches. For example:

Display brief MSDP peer information on Switch A.

```
[SwitchA] display msdp brief
```

```
MSDP Peer Brief Information of VPN-Instance: public net
```

Configured	Up	Listen	Connect	Shutdown	Down
2	2	0	0	0	0

Peer's Address	State	Up/Down time	AS	SA Count	Reset Count
10.110.3.2	Up	01:07:08	?	8	0
10.110.6.2	Up	00:16:39	?	13	0

Display brief MSDP peer information on Switch D.

```
[SwitchD] display msdp brief
```

```
MSDP Peer Brief Information of VPN-Instance: public net
```

Configured	Up	Listen	Connect	Shutdown	Down
1	1	0	0	0	0

Peer's Address	State	Up/Down time	AS	SA Count	Reset Count
10.110.1.1	Up	01:07:09	?	8	0

Display brief MSDP peer information on Switch G.

```
[SwitchG] display msdp brief
```

```
MSDP Peer Brief Information of VPN-Instance: public net
```

Configured	Up	Listen	Connect	Shutdown	Down
1	1	0	0	0	0

Peer's Address	State	Up/Down time	AS	SA Count	Reset Count
10.110.2.1	Up	00:16:40	?	13	0

Anycast RP configuration

Network requirements

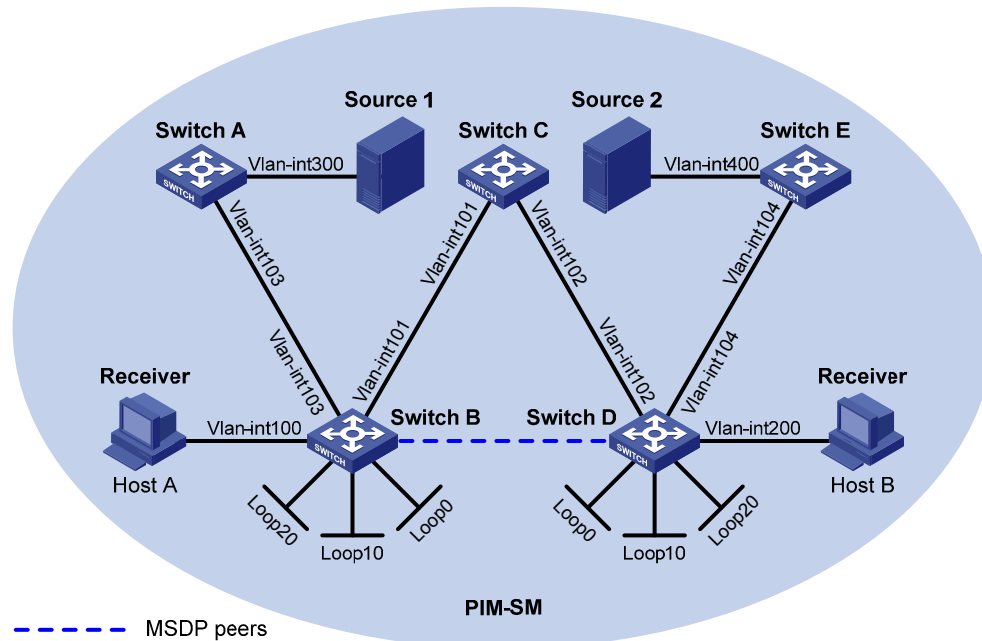
As shown in [Figure 62](#), the PIM-SM domain has multiple multicast sources and receivers. OSPF runs within the domain to provide unicast routes.

Configure the Anycast RP application so that the receiver-side DRs and the source-side DRs can initiate a join message to their respective RPs that are the topologically nearest to them.

On Switch B and Switch D, configure the interface Loopback 10 as a C-BSR, and Loopback 20 as a C-RP.

The router ID of Switch B is 1.1.1.1, and the router ID of Switch D is 2.2.2.2. Set up an MSDP peering relationship between Switch B and Switch D.

Figure 62 Network diagram



Device	Interface	IP address	Device	Interface	IP address
Source 1	—	10.110.5.100/24	Switch C	Vlan-int101	192.168.1.2/24
Source 2	—	10.110.6.100/24		Vlan-int102	192.168.2.2/24
Switch A	Vlan-int300	10.110.5.1/24	Switch D	Vlan-int200	10.110.3.1/24
	Vlan-int103	10.110.2.2/24		Vlan-int104	10.110.4.1/24
Switch B	Vlan-int100	10.110.1.1/24		Vlan-int102	192.168.2.1/24
	Vlan-int103	10.110.2.1/24		Loop0	2.2.2.2/32
	Vlan-int101	192.168.1.1/24		Loop10	4.4.4.4/32
	Loop0	1.1.1.1/32		Loop20	10.1.1.1/32
	Loop10	3.3.3.3/32	Switch E	Vlan-int400	10.110.6.1/24
	Loop20	10.1.1.1/32		Vlan-int104	10.110.4.2/24

Configuration procedure

1. Configure IP addresses and unicast routing:
Configure the IP address and subnet mask for each interface as per Figure 62. (Details not shown.)
Configure OSPF for interconnection between the switches. Ensure the network-layer interoperation among the switches, and ensure the dynamic update of routing information between the switches through a unicast routing protocol. (Details not shown.)

2. Enable IP multicast routing, and enable PIM-SM and IGMP:

Enable IP multicast routing on Switch B, enable PIM-SM on each interface, and enable IGMP on the host-side interface VLAN-interface 100.

```
<SwitchB> system-view
[SwitchB] multicast routing-enable
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] igmp enable
[SwitchB-Vlan-interface100] pim sm
[SwitchB-Vlan-interface100] quit
```

```

[SwitchB] interface vlan-interface 103
[SwitchB-Vlan-interface103] pim sm
[SwitchB-Vlan-interface103] quit
[SwitchB] interface Vlan-interface 101
[SwitchB-Vlan-interface101] pim sm
[SwitchB-Vlan-interface101] quit
[SwitchB] interface loopback 0
[SwitchB-LoopBack0] pim sm
[SwitchB-LoopBack0] quit
[SwitchB] interface loopback 10
[SwitchB-LoopBack10] pim sm
[SwitchB-LoopBack10] quit
[SwitchB] interface loopback 20
[SwitchB-LoopBack20] pim sm
[SwitchB-LoopBack20] quit

```

The configuration on Switch A, Switch C, Switch D, and Switch E is similar to the configuration on Switch B.

3. Configure C-BSRs and C-RPs:

Configure Loopback 10 as a C-BSR and Loopback 20 as a C-RP on Switch B.

```

[SwitchB] pim
[SwitchB-pim] c-bsr loopback 10
[SwitchB-pim] c-rp loopback 20
[SwitchB-pim] quit

```

The configuration on Switch D is similar to the configuration on Switch B.

4. Configure MSDP peers:

Configure an MSDP peer on Loopback 0 of Switch B.

```

[SwitchB] msdp
[SwitchB-msdp] originating-rp loopback 0
[SwitchB-msdp] peer 2.2.2.2 connect-interface loopback 0
[SwitchB-msdp] quit

```

Configure an MSDP peer on Loopback 0 of Switch D.

```

[SwitchD] msdp
[SwitchD-msdp] originating-rp loopback 0
[SwitchD-msdp] peer 1.1.1.1 connect-interface loopback 0
[SwitchD-msdp] quit

```

5. Verify the configuration:

You can use the **display msdp brief** command to view the brief information of MSDP peering relationships between the switches.

Display brief MSDP peer information on Switch B.

```
[SwitchB] display msdp brief
```

MSDP Peer Brief Information of VPN-Instance: public net

Configured	Up	Listen	Connect	Shutdown	Down
1	1	0	0	0	0

Peer's Address	State	Up/Down time	AS	SA Count	Reset Count
2.2.2.2	Up	00:10:17	?	0	0

Display brief MSDP peer information on Switch D.

```
[SwitchD] display msdp brief
MSDP Peer Brief Information of VPN-Instance: public net
  Configured   Up           Listen       Connect      Shutdown     Down
  1            1           0           0           0           0

  Peer's Address   State   Up/Down time   AS   SA Count   Reset Count
  1.1.1.1         Up     00:10:18      ?   0         0
```

To view the PIM routing information on the switches, use the **display pim routing-table** command. When Source 1 (10.110.5.100/24) sends multicast data to multicast group G (225.1.1.1), Host A joins multicast group G. By comparing the PIM routing information displayed on Switch B with that displayed on Switch D, you can see that Switch B acts now as the RP for Source 1 and Host A.

Display PIM routing information on Switch B.

```
[SwitchB] display pim routing-table
VPN-Instance: public net
Total 1 (*, G) entry; 1 (S, G) entry

(*, 225.1.1.1)
  RP: 10.1.1.1 (local)
  Protocol: pim-sm, Flag: WC
  UpTime: 00:15:04
  Upstream interface: Register
    Upstream neighbor: NULL
    RPF prime neighbor: NULL
  Downstream interface(s) information:
  Total number of downstreams: 1
    1: Vlan-interface100
      Protocol: igmp, UpTime: 00:15:04, Expires: -

(10.110.5.100, 225.1.1.1)
  RP: 10.1.1.1 (local)
  Protocol: pim-sm, Flag: SPT 2MSDP ACT
  UpTime: 00:46:28
  Upstream interface: Vlan-interface103
    Upstream neighbor: 10.110.2.2
    RPF prime neighbor: 10.110.2.2
  Downstream interface(s) information:
  Total number of downstreams: 1
    1: Vlan-interface100
      Protocol: pim-sm, UpTime: - , Expires: -
```

Display PIM routing information on Switch D.

```
[SwitchD] display pim routing-table
```

No information is output on Switch D.

Host A has left multicast group G. Source 1 has stopped sending multicast data to multicast group G. When Source 2 (10.110.6.100/24) sends multicast data to G, Host B joins G. By comparing the PIM routing information displayed on Switch B with that displayed on Switch D, you can see that Switch D acts now as the RP for Source 2 and Host B.

```

# Display PIM routing information on Switch B.
[SwitchB] display pim routing-table
No information is output on Switch B.
# Display PIM routing information on Switch D.
[SwitchD] display pim routing-table
VPN-Instance: public net
Total 1 (*, G) entry; 1 (S, G) entry

(*, 225.1.1.1)
  RP: 10.1.1.1 (local)
  Protocol: pim-sm, Flag: WC
  UpTime: 00:12:07
  Upstream interface: Register
    Upstream neighbor: NULL
    RPF prime neighbor: NULL
  Downstream interface(s) information:
  Total number of downstreams: 1
    1: Vlan-interface200
      Protocol: igmp, UpTime: 00:12:07, Expires: -

(10.110.6.100, 225.1.1.1)
  RP: 10.1.1.1 (local)
  Protocol: pim-sm, Flag: SPT 2MSDP ACT
  UpTime: 00:40:22
  Upstream interface: Vlan-interface104
    Upstream neighbor: 10.110.4.2
    RPF prime neighbor: 10.110.4.2
  Downstream interface(s) information:
  Total number of downstreams: 1
    1: Vlan-interface200
      Protocol: pim-sm, UpTime: - , Expires: -

```

SA message filtering configuration

Network requirements

As shown in [Figure 63](#), three PIM-SM domains exist in the network, and OSPF runs within and among the domains to provide unicast routing.

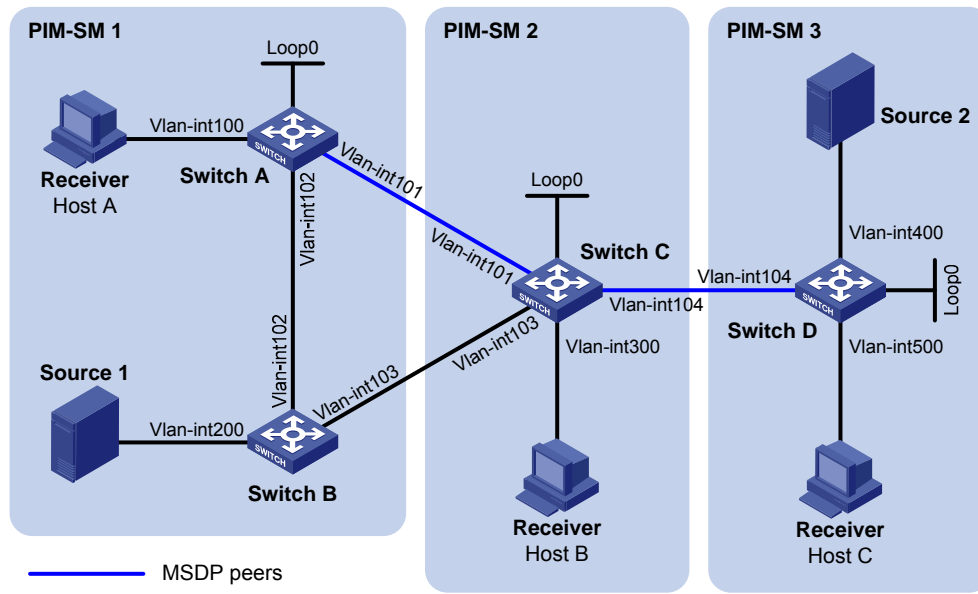
Loopback 0 is configured as a C-BSR and a C-RP in the related PIM-SM domains on Switch A, Switch C, and Switch D, respectively.

An MSDP peering relationship is set up between Switch A and Switch C and between Switch C and Switch D.

Source 1 sends multicast data to multicast groups 225.1.1.0/30 and 226.1.1.0/30, and Source 2 sends multicast data to multicast group 227.1.1.0/30.

Configure SA message filtering rules so that receivers Host A and Host B can receive only the multicast data addressed to multicast groups 225.1.1.0/30 and 226.1.1.0/30, and Host can receive only the multicast data addressed to multicast groups 226.1.1.0/30 and 227.1.1.0/30.

Figure 63 Network diagram



Device	Interface	IP address	Device	Interface	IP address
Source 1	—	10.110.3.100/24	Switch C	Vlan-int300	10.110.4.1/24
Source 2	—	10.110.6.100/24		Vlan-int104	10.110.5.1/24
Switch A	Vlan-int100	10.110.1.1/24		Vlan-int101	192.168.1.2/24
	Vlan-int102	10.110.2.1/24		Vlan-int103	192.168.2.2/24
	Vlan-int101	192.168.1.1/24		Loop0	2.2.2.2/32
	Loop0	1.1.1.1/32	Switch D	Vlan-int400	10.110.6.1/24
Switch B	Vlan-int200	10.110.3.1/24		Vlan-int500	10.110.7.1/24
	Vlan-int102	10.110.2.2/24		Vlan-int104	10.110.5.2/24
	Vlan-int103	192.168.2.1/24		Loop0	3.3.3.3/32

Configuration Procedure

- Configure IP addresses and unicast routing:
Configure the IP address and subnet mask for each interface as per Figure 63. (Details not shown.)
Configure OSPF for interoperation among the switches. Ensure the network-layer interoperation within and between the PIM-SM domains and ensure dynamic update of routing information among the switches by leveraging unicast routing. (Details not shown.)
- Enable IP multicast routing, PIM-SM and IGMP, and configure a PIM domain border:
On Switch A, enable IP multicast routing, enable PIM-SM on each interface, and enable IGMP on the host-side interface, VLAN-interface 100.

```

<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] igmp enable
[SwitchA-Vlan-interface100] pim sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim sm
[SwitchA-Vlan-interface101] quit

```



```
[SwitchA] interface vlan-interface 102
[SwitchA-Vlan-interface102] pim sm
[SwitchA-Vlan-interface102] quit
[SwitchA] interface loopback 0
[SwitchA-LoopBack0] pim sm
[SwitchA-LoopBack0] quit
```

The configuration on Switch B, Switch C and Switch D is similar to the configuration on Switch A. (Details not shown.)

Configure a PIM domain border on Switch C.

```
[SwitchC] interface vlan-interface 101
[SwitchC-Vlan-interface101] pim bsr-boundary
[SwitchC-Vlan-interface101] quit
[SwitchC] interface vlan-interface 103
[SwitchC-Vlan-interface103] pim bsr-boundary
[SwitchC-Vlan-interface103] quit
[SwitchC] interface vlan-interface 104
[SwitchC-Vlan-interface104] pim bsr-boundary
[SwitchC-Vlan-interface104] quit
```

The configuration on Switch A, Switch B and Switch D is similar to the configuration on Switch C. (Details not shown.)

3. Configure C-BSRs and C-RPs:

Configure Loopback 0 as a C-BSR and a C-RP on Switch A.

```
[SwitchA] pim
[SwitchA-pim] c-bsr loopback 0
[SwitchA-pim] c-rp loopback 0
[SwitchA-pim] quit
```

The configuration on Switch C and Switch D is similar to the configuration on Switch A. (Details not shown.)

4. Configure MSDP peers:

Configure an MSDP peer on Switch A.

```
[SwitchA] msdp
[SwitchA-msdp] peer 192.168.1.2 connect-interface vlan-interface 101
[SwitchA-msdp] quit
```

Configure MSDP peers on Switch C.

```
[SwitchC] msdp
[SwitchC-msdp] peer 192.168.1.1 connect-interface vlan-interface 101
[SwitchC-msdp] peer 10.110.5.2 connect-interface vlan-interface 104
[SwitchC-msdp] quit
```

Configure an MSDP peer on Switch D.

```
[SwitchD] msdp
[SwitchD-msdp] peer 10.110.5.1 connect-interface vlan-interface 104
[SwitchD-msdp] quit
```

5. Configure SA message filtering rules:

Configure an SA message rule on Switch C so that Switch C will not forward SA messages for (Source 1, 225.1.1.0/30) to Switch D.

```
[SwitchC] acl number 3001
```

```
[SwitchC-acl-adv-3001] rule deny ip source 10.110.3.100 0 destination 225.1.1.0
0.0.0.3
[SwitchC-acl-adv-3001] rule permit ip source any destination any
[SwitchC-acl-adv-3001] quit
[SwitchC] msdp
[SwitchC-msdp] peer 10.110.5.2 sa-policy export acl 3001
[SwitchC-msdp] quit
```

Configure an SA message rule on Switch D so that Switch D will not create SA messages for Source 2.

```
[SwitchD] acl number 2001
[SwitchD-acl-basic-2001] rule deny source 10.110.6.100 0
[SwitchD-acl-basic-2001] quit
[SwitchD] msdp
[SwitchD-msdp] import-source acl 2001
[SwitchD-msdp] quit
```

6. Verify the configuration:

View the (S, G) entries cached in the SA cache on the switches using the **display msdp sa-cache** command. For example:

Display the (S, G) entries cached in the SA cache on Switch C.

```
[SwitchC] display msdp sa-cache
MSDP Source-Active Cache Information of VPN-Instance: public net
MSDP Total Source-Active Cache - 8 entries
MSDP matched 8 entries
```

(Source, Group)	Origin RP	Pro	AS	Uptime	Expires
(10.110.3.100, 225.1.1.0)	1.1.1.1	?	?	02:03:30	00:05:31
(10.110.3.100, 225.1.1.1)	1.1.1.1	?	?	02:03:30	00:05:31
(10.110.3.100, 225.1.1.2)	1.1.1.1	?	?	02:03:30	00:05:31
(10.110.3.100, 225.1.1.3)	1.1.1.1	?	?	02:03:30	00:05:31
(10.110.3.100, 226.1.1.0)	1.1.1.1	?	?	02:03:30	00:05:31
(10.110.3.100, 226.1.1.1)	1.1.1.1	?	?	02:03:30	00:05:31
(10.110.3.100, 226.1.1.2)	1.1.1.1	?	?	02:03:30	00:05:31
(10.110.3.100, 226.1.1.3)	1.1.1.1	?	?	02:03:30	00:05:31

Display the (S, G) entries cached in the SA cache on Switch D.

```
[SwitchD] display msdp sa-cache
MSDP Source-Active Cache Information of VPN-Instance: public net
MSDP Total Source-Active Cache - 4 entries
MSDP matched 4 entries
```

(Source, Group)	Origin RP	Pro	AS	Uptime	Expires
(10.110.3.100, 226.1.1.0)	1.1.1.1	?	?	00:32:53	00:05:07
(10.110.3.100, 226.1.1.1)	1.1.1.1	?	?	00:32:53	00:05:07
(10.110.3.100, 226.1.1.2)	1.1.1.1	?	?	00:32:53	00:05:07
(10.110.3.100, 226.1.1.3)	1.1.1.1	?	?	00:32:53	00:05:07

Troubleshooting MSDP

MSDP peers stay in down state

Symptom

The configured MSDP peers stay in the down state.

Analysis

- A TCP connection-based MSDP peering relationship is established between the local interface address and the MSDP peer after the configuration.
- The TCP connection setup will fail if the local interface address is not consistent with the MSDP peer address configured on the peer router.
- If no route is available between the MSDP peers, the TCP connection setup will fail.

Solution

1. Verify that a route is available between the routers. Use the **display ip routing-table** command to determine whether the unicast route between the routers is correct.
2. Verify that a unicast route is available between the two routers that will become MSDP peers to each other.
3. Verify the interface address consistency between the MSDP peers. Use the **display current-configuration** command to verify that the local interface address and the MSDP peer address of the remote router are the same.

No SA entries in the switch's SA cache

Symptom

MSDP fails to send (S, G) entries through SA messages.

Analysis

- The **import-source** command controls sending (S, G) entries through SA messages to MSDP peers. If this command is executed without the *acl-number* argument, all the (S, G) entries are filtered out. That is, no (S, G) entries of the local domain is advertised.
- If the **import-source** command is not executed, the system will advertise all the (S, G) entries of the local domain. If MSDP fails to send (S, G) entries through SA messages, verify that the **import-source** command has been correctly configured.

Solution

1. Use the **display ip routing-table** command to verify that the unicast route between the routers is correct.
2. Verify that a unicast route is available between the two routers that will become MSDP peers to each other.
3. Verify the configuration of the **import-source** command and its *acl-number* argument and be sure that ACL rule can filter appropriate (S, G) entries.

Inter-RP communication faults in Anycast RP application

Symptom

RPs fail to exchange their locally registered (S, G) entries with one another in the Anycast RP application.

Analysis

- In the Anycast RP application, RPs in the same PIM-SM domain are configured to be MSDP peers to achieve load balancing among the RPs.
- An MSDP peer address must be different from the Anycast RP address, and the C-BSR and C-RP must be configured on different devices or interfaces.
- If the **originating-rp** command is executed, MSDP will replace the RP address in the SA messages with the address of the interface specified in the command.
- When an MSDP peer receives an SA message, it performs RPF check on the message. If the MSDP peer finds that the remote RP address is the same as the local RP address, it will discard the SA message.

Solution

1. Use the **display ip routing-table** command to verify that the unicast route between the routers is correct.
2. Verify that a unicast route is available between the two routers that will become MSDP peer to each other.
3. Verify the configuration of the **originating-rp** command. In the Anycast RP application environment, be sure to use the **originating-rp** command to configure the RP address in the SA messages, which must be the local interface address.
4. Verify that the C-BSR address is different from the Anycast RP address.

Configuring MBGP (available only on the HP 5500 EI)

MBGP overview

This chapter covers configuration tasks related to multiprotocol BGP for IP multicast only. For more information about BGP, see *Layer 3—IP Routing Configuration Guide*. The term "router" in this chapter refers to both routers and Layer 3 switches.

BGP-4 can carry routing information for IPv4 only. IETF defined Multiprotocol Border Gateway Protocol (MP-BGP) to extend BGP-4 so that BGP can carry routing information for multiple network-layer protocols.

For a network, the topology for multicast might be different from that for unicast. To distinguish them, the MP-BGP enables BGP to carry the unicast Network Layer Reachability Information (NLRI) and multicast NLRI separately. The multicast NLRI performs reverse path forwarding (RPF) exclusively. In this way, route selection for a destination through the unicast routing table and through the multicast routing table have different results, ensuring consistent unicast forwarding and normal multicast between domains. For more information about RPF, see "[Configuring multicast routing and forwarding \(available only on the HP 5500 EI\)](#)."

MP-BGP is defined in RFC 2858 (Multiprotocol Extensions for BGP-4). The application of MP-BGP on multicast is called Multicast BGP (MBGP).

Protocols and standards

- RFC 2858, *Multiprotocol Extensions for BGP-4*
- RFC 3392, *Capabilities Advertisement with BGP-4*
- draft-ietf-idmr-bgp-mcast-attr-00, *BGP Attributes for Multicast Tree Construction*
- RFC 4271, *A Border Gateway Protocol 4 (BGP-4)*
- RFC 5291, *Outbound Route Filtering Capability for BGP-4*
- RFC 5292, *Address-Prefix-Based Outbound Route Filter for BGP-4*

MBGP configuration task list

Task	Remarks	
Configuring basic MBGP functions	Required	
Configuring MBGP route redistribution	Required	
Controlling route advertisement and reception	Configuring default route redistribution into MBGP	Optional
	Configuring MBGP route summarization	Optional
	Advertising a default route to an IPv4 MBGP peer or peer group	Optional
	Configuring outbound MBGP route filtering	Optional

Task		Remarks
	Configuring inbound MBGP route filtering	Optional
	Configuring MBGP route dampening	Optional
Configuring MBGP route attributes	Configuring MBGP route preferences	
	Configuring the default local preference	
	Configuring the MED attribute	Optional
	Configuring the next hop attribute	
	Configuring the AS-PATH attributes	
Tuning and optimizing MBGP networks	Configuring MBGP soft reset	Optional
	Enabling the MBGP ORF capability	Optional
	Configuring the maximum number of MBGP routes for load balancing	Optional
Configuring a large scale MBGP network	Configuring IPv4 MBGP peer groups	Optional
	Configuring MBGP community	Optional
	Configuring an MBGP route reflector	Optional

Configuring basic MBGP functions

Configuration prerequisites

Before you configure MBGP, be sure that neighboring nodes can access each other at the network layer.

Configuration procedure

To configure basic MBGP functions:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Specify a peer or peer group and its AS number.	peer { <i>group-name</i> <i>ip-address</i> } as-number <i>as-number</i>	Not specified by default.
4. Enter IPv4 MBGP address family view.	ipv4-family multicast	N/A
5. Enable a peer or peer group created in IPv4 unicast view.	peer { <i>group-name</i> <i>ip-address</i> } enable	Not enabled by default.
6. Specify a preferred value for routes from an IPv4 MBGP peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> } preferred-value <i>value</i>	Optional. The default preferred value is 0.

Controlling route advertisement and reception

Configuration prerequisites

You need to configure basic MBGP functions before configuring this task.

Configuring MBGP route redistribution

MBGP can advertise routing information in the local AS to neighboring ASs. It redistributes such routing information from IGP into its routing table rather than learning the information by itself.

To configure MBGP route redistribution:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv4 MBGP address family view.	ipv4-family multicast	N/A
4. Configure MBGP route redistribution.	<p>Enable route redistribution from another routing protocol:</p> <pre>import-route <i>protocol</i> [{ <i>process-id</i> all-processes } [allow-direct med <i>med-value</i> route-policy <i>route-policy-name</i>] *]</pre> <p>Inject a network into the MBGP routing table:</p> <pre>network <i>ip-address</i> [<i>mask</i> <i>mask-length</i>] [short-cut route-policy <i>route-policy-name</i>]</pre>	<p>Use either command</p> <p>No route redistribution is configured by default.</p> <p>The allow-direct keyword is available only when the specified routing protocol is OSPF.</p> <p>The Origin attribute of routes redistributed into the MBGP routing table with the import-route command is Incomplete.</p> <p>The Origin attribute of routes injected into the MBGP routing table with the network command is IGP.</p>

NOTE:

The networks to be injected must exist in the local IP routing table, and using a routing policy makes route control more flexible.

Configuring default route redistribution into MBGP

You cannot use the **import-route** command to redistribute any default route into the MBGP routing table. This task allows you to redistribute default routes in another way.

To configure MBGP to redistribute a default route from another protocol:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A

Step	Command	Remarks
3. Enter MBGP address family view.	ipv4-family multicast	N/A
4. Enable route redistribution from another routing protocol.	import-route <i>protocol</i> [{ <i>process-id</i> all-processes } [allow-direct med <i>med-value</i> route-policy <i>route-policy-name</i>] *]	No route redistribution is configured by default. The allow-direct keyword is available only when the specified routing protocol is OSPF.
5. Enable default route redistribution into the MBGP routing table.	default-route imported	Not enabled by default.

Configuring MBGP route summarization

To reduce the routing table size on medium and large MBGP networks, you need to configure route summarization on peers. MBGP supports automatic and manual summarization modes:

- **Automatic summarization**—Summarizes subnets redistributed from IGP. With the feature configured, MBGP advertises only summary natural networks rather than subnets. The default routes and routes injected with the **network** command are not summarized.
- **Manual summarization**—Summarizes MBGP local routes. A manual summary route has a higher priority than an automatic one.

To configure MBGP route summarization:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv4 MBGP address family view.	ipv4-family multicast	N/A
4. Configure MBGP route summarization.	Enable automatic route summarization: summary automatic Configure manual route summarization: aggregate <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } [as-set attribute-policy <i>route-policy-name</i> detail-suppressed origin-policy <i>route-policy-name</i> suppress-policy <i>route-policy-name</i>] *	Use either command No route summarization is configured by default. If the two commands are both configured, the manual route summarization takes effect.

Advertising a default route to an IPv4 MBGP peer or peer group

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv4 MBGP address family view.	ipv4-family multicast	N/A
4. Advertise a default route to an MBGP peer or peer group.	peer { <i>group-name</i> <i>ip-address</i> } default-route-advertise [route-policy <i>route-policy-name</i>]	Not advertised by default.

NOTE:

With the **peer default-route-advertise** command executed, the router sends a default route with the next hop as itself to the specified MBGP peer or peer group, whether the default route is available or not in the routing table.

Configuring outbound MBGP route filtering

If several filtering policies are configured, they are applied in the following sequence:

1. **filter-policy export**
2. **peer filter-policy export**
3. **peer as-path-acl export**
4. **peer ip-prefix export**
5. **peer route-policy export**

Only the routes that have passed all the configured policies can be advertised.

To configure BGP route distribution filtering policies:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv4 MBGP address family view.	ipv4-family multicast	N/A

Step	Command	Remarks
4. Configure BGP route distribution filtering policies.	<ul style="list-style-type: none"> Configure the filtering of redistributed routes: filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } export [direct isis <i>process-id</i> ospf <i>process-id</i> rip <i>process-id</i> static] Apply a routing policy to advertisements to an IPv4 MBGP peer or a peer group: peer { <i>group-name</i> <i>peer-address</i> } route-policy <i>route-policy-name</i> export Reference an ACL to filter advertisements to an IPv4 MBGP peer or a peer group: peer { <i>group-name</i> <i>ip-address</i> } filter-policy <i>acl-number</i> export Reference an AS path ACL to filter route advertisements to an IPv4 MBGP peer or a peer group: peer { <i>group-name</i> <i>ip-address</i> } as-path-acl <i>as-path-acl-number</i> export Reference an IP prefix list to filter route advertisements to an IPv4 MBGP peer or a peer group: peer { <i>group-name</i> <i>ip-address</i> } ip-prefix <i>ip-prefix-name</i> export 	Use one of these commands No outbound route filtering is configured by default.

Configuring inbound MBGP route filtering

By configuring MBGP route reception filtering policies, you can filter out unqualified routes from an MBGP peer or peer group. Members of a peer group can have different route reception filtering policies from the peer group.

If several filtering policies are configured, they are applied in the following sequence:

1. **filter-policy import**
2. **peer filter-policy import**
3. **peer as-path-acl import**
4. **peer ip-prefix import**
5. **peer route-policy import**

Only the routes that have passed all the configured policies can be advertised.

To configure MBGP route reception filtering policies:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv4 MBGP address family view.	ipv4-family multicast	N/A

Step	Command	Remarks
4. Configure MBGP route reception filtering policies.	<ul style="list-style-type: none"> Filter incoming routes using an ACL or IP prefix list: filter-policy { <i>acl-number</i> ip-prefix <i>ip-prefix-name</i> } import Reference a routing policy to routes from an IPv4 MBGP peer or a peer group: peer { <i>group-name</i> <i>ip-address</i> } route-policy <i>policy-name</i> import Reference an ACL to filter routing information from an IPv4 MBGP peer or a peer group: peer { <i>group-name</i> <i>ip-address</i> } filter-policy <i>acl-number</i> import Reference an AS path ACL to filter routing information from an IPv4 MBGP peer or a peer group: peer { <i>group-name</i> <i>ip-address</i> } as-path-acl <i>as-path-acl-number</i> import Reference an IP prefix list to filter routing information from an IPv4 MBGP peer or a peer group: peer { <i>group-name</i> <i>ip-address</i> } ip-prefix <i>ip-prefix-name</i> import 	Use one of these command No inbound route filtering is configured by default.
5. Specify the maximum number of routes that can be received from an IPv4 MBGP peer or a peer group.	peer { <i>group-name</i> <i>ip-address</i> } route-limit <i>limit</i> [<i>percentage</i>]	Optional The number is unlimited by default.

Configuring MBGP route dampening

By configuring MBGP route dampening, you can suppress unstable routes from being added to the MBGP routing table or being advertised to MBGP peers.

To configure BGP route dampening:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv4 MBGP address family view.	ipv4-family multicast	N/A

Step	Command	Remarks
4. Configure BGP route dampening parameters.	dampening [<i>half-life-reachable</i> <i>half-life-unreachable reuse</i> <i>suppress ceiling</i> route-policy <i>route-policy-name</i>] *	Not configured by default.

Configuring MBGP route attributes

You can modify MBGP route attributes to affect route selection.

Configuration prerequisites

Before you configure this task, you need to configure basic MBGP functions.

Configuring MBGP route preferences

You can reference a routing policy to set preferences for routes matching it. Routes not matching it use the default preferences.

To configure MBGP route preferences:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv4 MBGP address family view.	ipv4-family multicast	N/A
4. Configure preferences for external, internal, and local MBGP routes.	preference { <i>external-preference</i> <i>internal-preference</i> <i>local-preference</i> route-policy <i>route-policy-name</i> }	Optional. The default preferences of multicast MBGP EBGP, MBGP IBGP, and local MBGP routes are 255, 255, and 130, respectively.

Configuring the default local preference

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv4 MBGP address family view.	ipv4-family multicast	N/A
4. Configure the default local preference.	default local-preference <i>value</i>	Optional. 100 by default.

Configuring the MED attribute

When other conditions of routes to a destination are identical, the route with the smallest MED is selected.

To configure the MED attribute:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv4 MBGP address family view.	ipv4-family multicast	N/A
4. Configure the default MED value.	default med <i>med-value</i>	Optional. 0 by default.
5. Enable the comparison of the MED of routes from different ASs.	compare-different-as-med	Optional. Not enabled by default.
6. Enable the comparison of the MED of routes from each AS.	bestroute compare-med	Optional. Not enabled by default.
7. Enable the comparison of the MED of routes from confederation peers.	bestroute med-confederation	Optional. Not enabled by default.

Configuring the next hop attribute

You can use the **peer next-hop-local** command to specify the local router as the next hop of routes sent to an MBGP IBGP peer or peer group. If load balancing is configured, the router specifies itself as the next hop of route advertisements to the multicast IBGP peer or peer group regardless of whether the **peer next-hop-local** command is configured.

In a third-party next-hop network (that is, when the local router has two multicast EBGP peers in a broadcast network), the router does not specify itself as the next hop of routing information sent to the EBGP peers unless the **peer next-hop-local** command is configured.

To specify the router as the next hop of routes sent to a peer or a peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv4 MBGP address family view.	ipv4-family multicast	N/A
4. Specify the router as the next hop of routes sent to a peer or a peer group.	peer { <i>group-name</i> <i>ip-address</i> } next-hop-local	Optional. By default, IPv6 MBGP specifies the local router as the next hop for routes sent to an EBGP peer or a peer group, but not for routes sent to an MBGP IBGP peer or a peer group.

Configuring the AS-PATH attributes

In general, MBGP checks whether the AS-PATH attribute of a route from a peer contains the local AS number. If it does, it discards the route to avoid routing loops.

To configure the AS-PATH attributes:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv4 MBGP address family view.	ipv4-family multicast	N/A
4. Allow the local AS number to appear in the AS-PATH of routes from a peer or a peer group and specify the number of times that the local AS number can appear in the AS-PATH of routes from the peer or the peer group.	peer { <i>group-name</i> <i>ip-address</i> } allow-as-loop [<i>number</i>]	Optional. By default, the local AS number cannot appear in routes from a peer or a peer group.
5. Disable BGP from considering the AS_PATH during best route selection.	bestroute as-path-neglect	Optional. By default, BGP considers AS_PATH during best route selection.
6. Configure updates to a peer or a peer group not to keep private AS numbers.	peer { <i>group-name</i> <i>ip-address</i> } public-as-only	Optional. By default, BGP updates carry private AS numbers.

Tuning and optimizing MBGP networks

This task involves resetting MBGP connections and configuring load balancing.

Configuration prerequisites

You need to configure basic MBGP functions before configuring this task.

Configuring MBGP soft reset

After modifying a route selection policy, you have to reset MBGP connections to make it take effect.

The current MBGP implementation supports the route refresh feature that enables dynamic route refresh without terminating MBGP connections.

However, if a peer that does not support route refresh exists in the network, you must configure the **peer keep-all-routes** command to save all routes from the peer. When the routing policy is changed, the system updates the MBGP routing table and applies the new policy.

Performing soft reset through route refresh

If the peer is enabled with route refresh, when the MBGP route selection policy is modified on a router, the router advertises a route-refresh message to its MBGP peers, which resend their routing information

to the router after receiving the message. Therefore, the local router can perform dynamic route update and apply the new policy without terminating MBGP connections.

To perform soft reset through route refresh:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A
3. Enable BGP route refresh for a peer or a peer group.	peer { group-name ip-address } capability-advertise route-refresh	Optional. Enabled by default.

Performing soft reset manually

If the peer does not support route refresh, you can use the **peer keep-all-routes** command to save all the route updates from the peer, and then use the **refresh bgp ipv4 multicast** command to soft-reset MBGP connections to refresh the MBGP routing table and apply the new policy without terminating MBGP connections.

To perform a manual soft reset:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A
3. Disable BGP route refresh and multiprotocol extensions for a peer or a peer group.	peer { group-name ip-address } capability-advertise conventional	Optional. Enabled by default.
4. Enter IPv4 MBGP address family view.	ipv4-family multicast	N/A
5. Keep all original routes from a peer or a peer group regardless of whether they pass the inbound filtering policies.	peer { group-name ip-address } keep-all-routes	Not kept by default.
6. Return to user view.	return	N/A
7. Soft-reset MBGP connections manually.	refresh bgp ipv4 multicast { all ip-address group group-name external internal } { export import }	Optional.

Enabling the MBGP ORF capability

The MBGP Outbound Router Filter (ORF) feature enables an MBGP speaker to send a set of ORFs to its MBGP peer through route-refresh messages. The peer then applies the ORFs, in addition to its local routing policies (if any), to filter updates to the MBGP speaker, reducing update messages and saving network resources.

After you enable the BGP ORF capability, the local BGP router negotiates the ORF capability with the BGP peer through Open messages. That is, the BGP router determines whether to carry ORF information in messages and, if so, whether to carry nonstandard ORF information in the packets. After completing

the negotiation process and establishing the neighboring relationship, the BGP router and its BGP peer can exchange ORF information through specific route-refresh messages.

For the parameters configured on both sides for ORF capability negotiation, see [Table 7](#).

To enable the MBGP ORF capability:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A
3. Enable BGP route refresh for a peer or a peer group.	peer { group-name ip-address } capability-advertise route-refresh	Optional. Enabled by default. If this feature is not enabled, you need to configure this command. For more information about the command, see <i>Layer 3—IP Routing Command Reference</i> .
4. Enable the non-standard BGP ORF capability for a BGP peer or a peer group.	peer { group-name ipv6-address } capability-advertise orf non-standard	Optional. By default, standard BGP ORF capability defined in RFC 5291 and RFC 5292 is supported. If this feature is not enabled, you need to configure this command. For more information about the command, see <i>Layer 3—IP Routing Command Reference</i> .
5. Enter MBGP address family view.	ipv4-family multicast	N/A
6. Enable the ORF IP prefix negotiation capability for an MBGP peer or a peer group.	peer { group-name ip-address } capability-advertise orf ip-prefix { both receive send }	Optional. Not enabled by default.

Table 7 Description of the both, send, and receive parameters and the negotiation result

Local parameter	Peer parameter	Negotiation result
send	<ul style="list-style-type: none"> • receive • both 	The ORF sending capability is enabled locally and the ORF receiving capability is enabled on the peer.
receive	<ul style="list-style-type: none"> • send • both 	The ORF receiving capability is enabled locally and the ORF sending capability is enabled on the peer.
both	both	Both the ORF sending and receiving capabilities are enabled locally and on the peer, respectively.

Configuring the maximum number of MBGP routes for load balancing

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv4 MBGP address family view.	ipv4-family multicast	N/A
4. Configure the maximum number of MBGP routes for load balancing.	balance <i>number</i>	Not configured by default.

Configuring a large scale MBGP network

Configuration prerequisites

Before you configure this task, you must make peering nodes accessible to each other at the network layer.

Configuring IPv4 MBGP peer groups

In a large-scale network, configuration and maintenance become difficult because of large numbers of MBGP peers. You can configure peer groups to make management easier and improve route distribution efficiency.

ⓘ IMPORTANT:

- To configure an MBGP peer group, you must enable the corresponding IPv4 BGP unicast peer group in IPv4 MBGP address family view.
- Before adding an MBGP peer to an MBGP peer group, you must add the corresponding IPv4 unicast peer to the IPv4 BGP peer group.

To configure an IPv4 MBGP peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Create a BGP peer group.	group <i>group-name</i> [external internal]	Not created by default.
4. Add a peer into the peer group.	peer <i>ip-address</i> group <i>group-name</i> [as-number <i>as-number</i>]	No peer is added by default.
5. Enter IPv4 MBGP address family view.	ipv4-family multicast	N/A
6. Enable the IPv4 unicast peer group.	peer <i>group-name</i> enable	N/A
7. Add an IPv4 MBGP peer to the peer group.	peer <i>ip-address</i> group <i>group-name</i>	Not configured by default.

Configuring MBGP community

The community attribute can be advertised between MBGP peers in different ASs. Routers in the same community share the same policy.

You can reference a routing policy to modify the community attribute for routes sent to a peer. In addition, you can define extended community attributes as needed.

When you configure MBGP community, you must reference a routing policy to define the specific community attributes, and apply the routing policy for route advertisement. For routing policy configuration, see *Layer 3—IP Routing Configuration Guide*.

To configure MBGP community:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A
3. Enter IPv4 MBGP address family view.	ipv4-family multicast	N/A
4. Advertise the community attribute to an MBGP peer or a peer group.	<ul style="list-style-type: none">Advertise the community attribute to an MBGP peer or a peer group: peer { group-name ip-address } advertise-communityAdvertise the extended community attribute to an MBGP peer or a peer group: peer { group-name ip-address } advertise-ext-community	<ul style="list-style-type: none">Use either commandNot configured by default.
5. Apply a routing policy to routes advertised to an MBGP peer or a peer group.	peer { group-name ip-address } route-policy route-policy-name export	Not configured by default.

Configuring an MBGP route reflector

To guarantee the connectivity between multicast IBGP peers in an AS, you need to make them fully meshed. But this becomes unpractical when large numbers of multicast IBGP peers exist. Configuring route reflectors can solve this problem.

In general, it is not required that clients of a route reflector be fully meshed. The route reflector forwards routing information between clients. If clients are fully meshed, you can disable route reflection between clients to reduce routing costs.

In general, a cluster has only one route reflector, and the router ID of the route reflector identifies the cluster. You can configure multiple route reflectors to improve network stability. In this case, you need to specify the same cluster ID for these route reflectors to avoid routing loops.

To configure an MBGP route reflector:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv4 MBGP address family view.	ipv4-family multicast	N/A
4. Configure the router as a route reflector and specify an MBGP peer or a peer group as its client.	peer { <i>group-name</i> <i>peer-address</i> } reflect-client	Not configured by default.
5. Enable route reflection between clients.	reflect between-clients	Optional. Enabled by default.
6. Configure the cluster ID of the route reflector.	reflector cluster-id <i>cluster-id</i>	Optional. By default, a route reflector uses its router ID as the cluster ID.

Displaying and maintaining MBGP

Displaying MBGP

Task	Command	Remarks
Display the IPv4 MBGP routing table.	display ip multicast routing-table [<i>verbose</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the IPv4 MBGP routing information matching the specified destination IP address.	display ip multicast routing-table <i>ip-address</i> [<i>mask-length</i> <i>mask</i>] [longer-match] [<i>verbose</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display MBGP peer group information.	display bgp multicast group [<i>group-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the advertised networks.	display bgp multicast network [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display AS path information.	display bgp multicast paths [<i>as-regular-expression</i> { begin exclude include } <i>regular-expression</i>]	Available in any view
Display MBGP peer information or peer group information.	display bgp multicast peer [[<i>ip-address</i>] <i>verbose</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the prefix entries in the ORF information from the specified BGP peer.	display bgp multicast peer <i>ip-address</i> received ip-prefix [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display MBGP routing information.	display bgp multicast routing-table [<i>ip-address</i> [{ <i>mask</i> <i>mask-length</i> } [longer-prefixes]]] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display MBGP routing information matching the AS path ACL.	display bgp multicast routing-table as-path-acl <i>as-path-acl-number</i> [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display MBGP CIDR routing information.	display bgp multicast routing-table cidr [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display MBGP routing information matching the specified BGP community.	display bgp multicast routing-table community [aa:nn&<1-13>] [no-advertise no-export no-export-subconfed] * [whole-match] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display MBGP routing information matching an MBGP community list.	display bgp multicast routing-table community-list { { <i>basic-community-list-number</i> <i>comm-list-name</i> } [whole-match] <i>adv-community-list-number</i> } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display MBGP dampened routing information.	display bgp multicast routing-table dampened [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display MBGP dampening parameter information.	display bgp multicast routing-table dampening parameter [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display MBGP routing information originating from different ASs.	display bgp multicast routing-table different-origin-as [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv4 MBGP routing flap statistics.	display bgp multicast routing-table flap-info [<i>regular-expression as-regular-expression</i> [as-path-acl <i>as-path-acl-number</i> <i>ip-address</i> [{ <i>mask</i> <i>mask-length</i> } [longer-match]]] [{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display IPv4 MBGP routing information sent to or received from an MBGP peer.	display bgp multicast routing-table peer <i>ip-address</i> { advertised-routes received-routes } [<i>network-address</i> [<i>mask</i> <i>mask-length</i>] statistic] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv4 MBGP routing information matching an AS regular expression.	display bgp multicast routing-table regular-expression <i>as-regular-expression</i>	Available in any view
Display IPv4 MBGP routing statistics.	display bgp multicast routing-table statistic [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Resetting MBGP connections

Task	Command	Remarks
Reset specified MBGP connections.	reset bgp ipv4 multicast { all <i>as-number</i> <i>ip-address</i> group <i>group-name</i> external internal }	Available in user view

Clearing MBGP information

Task	Command	Remarks
Clear dampened routing information and release suppressed routes.	reset bgp ipv4 multicast dampening [<i>ip-address</i> [<i>mask</i> <i>mask-length</i>]]	Available in user view

Task	Command	Remarks
Clear MBGP route flap statistics.	<code>reset bgp ipv4 multicast flap-info [regexp as-path-regexp as-path-acl as-path-acl-number ip-address [mask mask-length]]</code>	Available in user view

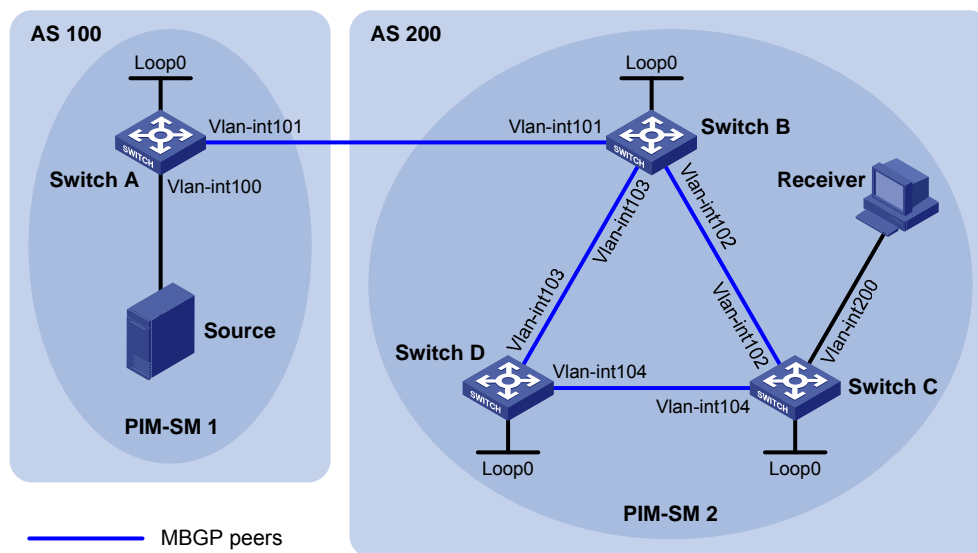
MBGP configuration example

Network requirements

As shown in the following figure:

- PIM-SM 1 is in AS 100, and PIM-SM 2 is in AS 200. OSPF is the IGP in the two ASs, and MBGP runs between the two ASs to exchange multicast route information.
- The multicast source belongs to PIM-SM 1, and the receiver belongs to PIM-SM 2.
- Configure the respective Loopback 0 of Switch A and Switch B as the C-BSR and C-RP of the respective PIM-SM domains.
- Set up an MSDP peer relationship between Switch A and Switch B through MBGP.

Figure 64 Network diagram



Device	Interface	IP address	Device	Interface	IP address
Source	N/A	10.110.1.100/24	Switch C	Vlan-int200	10.110.2.1/24
Switch A	Vlan-int100	10.110.1.1/24	Switch C	Vlan-int102	192.168.2.2/24
	Vlan-int101	192.168.1.1/24		Vlan-int104	192.168.4.1/24
	Loop0	1.1.1.1/32		Loop0	3.3.3.3/32
Switch B	Vlan-int101	192.168.1.2/24	Switch D	Vlan-int103	192.168.3.2/24
	Vlan-int102	192.168.2.1/24		Vlan-int104	192.168.4.2/24
	Vlan-int103	192.168.3.1/24		Loop0	4.4.4.4/32
	Loop0	2.2.2.2/32			

Configuration procedure

1. Configure IP addresses for interfaces as shown in Figure 64. (Details not shown.)
2. Configure OSPF. (Details not shown.)

3. Enable IP multicast routing, PIM-SM and IGMP, and configure a PIM-SM domain border:

Enable IP multicast routing on Switch A, and enable PIM-SM on each interface.

```
<SwitchA> system-view
[SwitchA] multicast routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] pim sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim sm
[SwitchA-Vlan-interface101] quit
```

The configuration on Switch B and Switch D is similar to the configuration on Switch A.

Enable IP multicast routing on Switch C, enable PIM-SM on each interface, and enable IGMP on the host-side interface VLAN-interface 200.

```
<SwitchC> system-view
[SwitchC] multicast routing-enable
[SwitchC] interface vlan-interface 102
[SwitchC-Vlan-interface102] pim sm
[SwitchC-Vlan-interface102] quit
[SwitchC] interface vlan-interface 104
[SwitchC-Vlan-interface104] pim sm
[SwitchC-Vlan-interface104] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] pim sm
[SwitchC-Vlan-interface200] igmp enable
[SwitchC-Vlan-interface200] quit
```

Configure a PIM domain border on Switch A.

```
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim bsr-boundary
[SwitchA-Vlan-interface101] quit
```

Configure a PIM domain border on Switch B.

```
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] pim bsr-boundary
[SwitchB-Vlan-interface101] quit
```

4. Configure Loopback 0 and the position of C-BSR, and C-RP:

Configure Loopback 0 and configure it as the C-BSR and C-RP on Switch A.

```
[SwitchA] interface loopback 0
[SwitchA-LoopBack0] ip address 1.1.1.1 32
[SwitchA-LoopBack0] pim sm
[SwitchA-LoopBack0] quit
[SwitchA] pim
[SwitchA-pim] c-bsr loopback 0
[SwitchA-pim] c-rp loopback 0
[SwitchA-pim] quit
```

Configure Loopback 0 and configure it as the C-BSR and C-RP on Switch B.

```
[SwitchB] interface loopback 0
[SwitchB-LoopBack0] ip address 2.2.2.2 32
[SwitchB-LoopBack0] pim sm
```

```
[SwitchB-LoopBack0] quit
[SwitchB] pim
[SwitchB-pim] c-bsr loopback 0
[SwitchB-pim] c-rp loopback 0
[SwitchB-pim] quit
```

5. Configure BGP, specify the MBGP peer and enable direct route redistribution:

On Switch A, configure the MBGP peer and enable direct route redistribution.

```
[SwitchA] bgp 100
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] peer 192.168.1.2 as-number 200
[SwitchA-bgp] import-route direct
[SwitchA-bgp] ipv4-family multicast
[SwitchA-bgp-af-mul] peer 192.168.1.2 enable
[SwitchA-bgp-af-mul] import-route direct
[SwitchA-bgp-af-mul] quit
[SwitchA-bgp] quit
```

On Switch B, configure the MBGP peer and enable route redistribution from OSPF.

```
[SwitchB] bgp 200
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] peer 192.168.1.1 as-number 100
[SwitchB-bgp] import-route ospf 1
[SwitchB-bgp] ipv4-family multicast
[SwitchB-bgp-af-mul] peer 192.168.1.1 enable
[SwitchB-bgp-af-mul] import-route ospf 1
[SwitchB-bgp-af-mul] quit
[SwitchB-bgp] quit
```

6. Configure MSDP peer:

Specify the MSDP peer on Switch A.

```
[SwitchA] msdp
[SwitchA-msdp] peer 192.168.1.2 connect-interface vlan-interface 101
[SwitchA-msdp] quit
```

Specify the MSDP peer on Switch B.

```
[SwitchB] msdp
[SwitchB-msdp] peer 192.168.1.1 connect-interface vlan-interface 101
[SwitchB-msdp] quit
```

7. Verify the configuration:

You can use the **display bgp multicast peer** command to display MBGP peers on a switch. For example, display MBGP peers on Switch B.

```
[SwitchB] display bgp multicast peer
```

```
BGP local router ID : 2.2.2.2
Local AS number : 200
Total number of peers : 3                Peers in established state : 3
```

Peer	AS	MsgRcvd	MsgSent	OutQ	PrefRcv	Up/Down	State
192.168.1.1	100	56	56	0	0	00:40:54	Established

You can use the **display msdp brief** command to display MSDP peers on a switch. For example, display brief information about MSDP peers on Switch B.

```
[SwitchB] display msdp brief
```

```
MSDP Peer Brief Information of VPN-Instance: public net
```

Configured	Up	Listen	Connect	Shutdown	Down
1	1	0	0	0	0
Peer's Address	State	Up/Down time	AS	SA Count	Reset Count
192.168.1.1	Up	00:07:17	100	1	0

Configuring MLD snooping

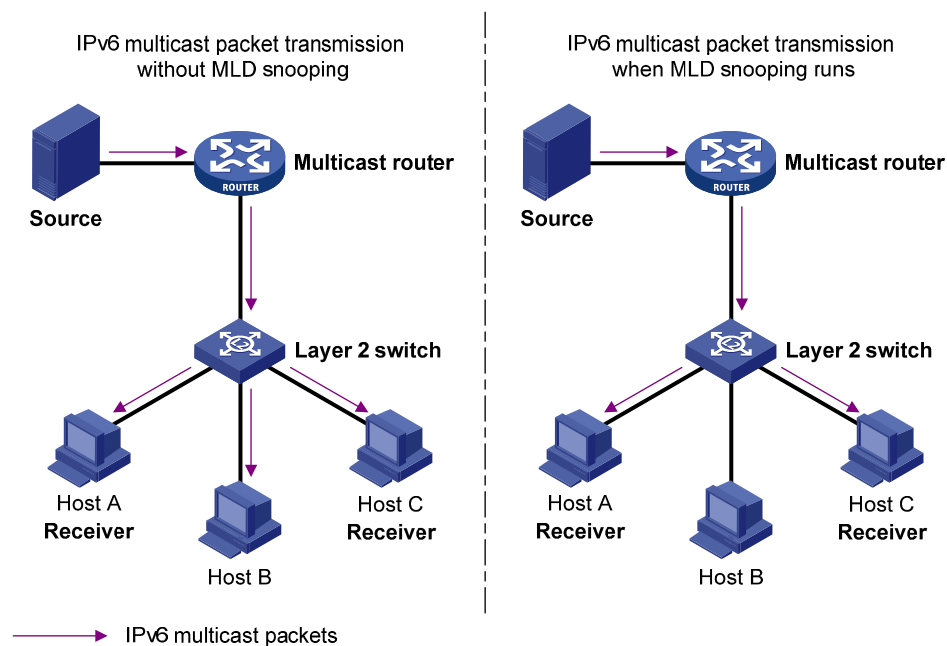
Overview

Multicast Listener Discovery (MLD) snooping is an IPv6 multicast constraining mechanism that runs on Layer 2 devices to manage and control IPv6 multicast groups.

By analyzing received MLD messages, a Layer 2 device that runs MLD snooping establishes mappings between ports and multicast MAC addresses and forwards IPv6 multicast data based on these mappings.

As shown in [Figure 65](#), when MLD snooping does not run on the Layer 2 switch, IPv6 multicast packets are flooded to all devices at Layer 2. When MLD snooping runs on the Layer 2 switch, multicast packets for known IPv6 multicast groups are multicast to the receivers, rather than flooded to all hosts at Layer 2.

Figure 65 Before and after MLD snooping is enabled on the Layer 2 device



MLD snooping enables the Layer 2 switch to forward IPv6 multicast data to only the receivers that require the data at Layer 2. It has the following advantages:

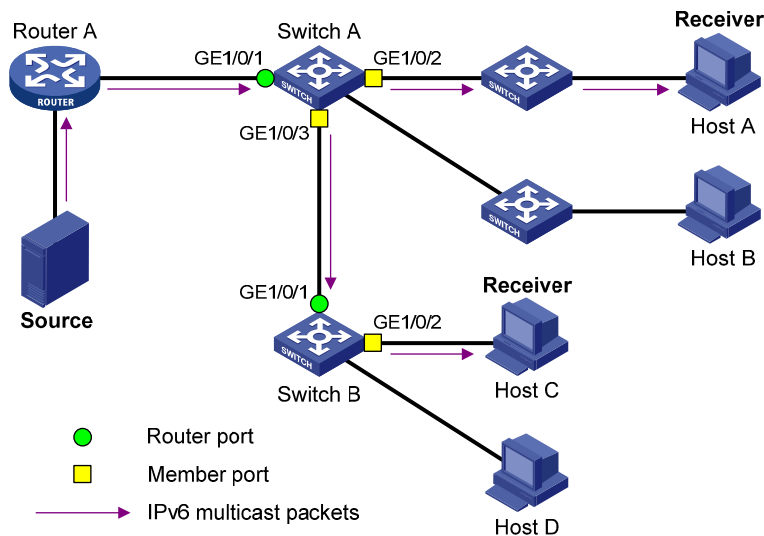
- Reducing Layer 2 broadcast packets, thus saving network bandwidth
- Enhancing the security of multicast traffic
- Facilitating the implementation of per-host accounting

Basic concepts in MLD snooping

MLD snooping related ports

As shown in [Figure 66](#), Router A connects to the multicast source, MLD snooping runs on Switch A and Switch B, and Host A and Host C are receiver hosts (namely, members of an IPv6 multicast group).

Figure 66 MLD snooping related ports



Ports involved in MLD snooping, as shown in Figure 66, are described as follows:

- Router port**—A router port is a port on the Ethernet switch that leads switch toward the Layer-3 multicast device (designated router or MLD querier). In the figure, GigabitEthernet 1/0/1 of Switch A and GigabitEthernet 1/0/1 of Switch B are router ports. The switch registers all its local router ports in its router port list.

In this document, a router port is port on a switch that leads the switch toward a Layer 3 multicast device. It is not a port on an ordinary router.

- Member port**—A member port (also known as IPv6 multicast group member port) is a port on the Ethernet switch that leads toward multicast group members. In the figure, GigabitEthernet 1/0/2 and GigabitEthernet 1/0/3 of Switch A and GigabitEthernet 1/0/2 of Switch B are member ports. The switch registers all the member ports on the local device in its MLD snooping forwarding table.

Unless otherwise specified, router ports and member ports in this document include both static and dynamic router ports and member ports.

NOTE:

An MLD snooping-enabled switch deems that the all its ports that receive MLD general queries with the source address other than 0::0 or that receive IPv6 PIM hello messages are dynamic router ports. For more information about IPv6 PIM hello messages, see "[Configuring IPv6 PIM \(available only on the HP 5500 EI\)](#)."

Aging timers for dynamic ports in MLD snooping and related messages and actions

Timer	Description	Message before expiry	Action after expiry
Dynamic router port aging timer	For each dynamic router port, the switch sets an aging timer. When the timer expires, the dynamic router port ages out.	MLD general query of which the source address is not 0::0 or IPv6 PIM hello.	The switch removes this port from its router port list.

Timer	Description	Message before expiry	Action after expiry
Dynamic member port aging timer	When a port dynamically joins a multicast group, the switch starts an aging timer for the port. When the timer expires, the dynamic member port ages out.	MLD report message.	The switch removes this port from the MLD snooping forwarding table.

NOTE:

In MLD snooping, only dynamic ports age out. Static ports never age out.

How MLD snooping works

In this section, the involved ports are dynamic ports. For information about how to configure and remove static ports, see "[Configuring static ports.](#)"

A switch that runs MLD snooping performs different actions when it receives different MLD messages, as follows:

When receiving a general query

The MLD querier periodically sends MLD general queries to all hosts and routers (FF02::1) on the local subnet to determine whether any active IPv6 multicast group members exist on the subnet.

After receiving an MLD general query, the switch forwards it to all ports in the VLAN, except the port that received the query. The switch also performs the following judgment:

- If the port that received the query is a dynamic router port in the router port list of the switch, the switch restarts the aging timer for the port.
- If the port is not in the router port list, the switch adds it into the router port list as a dynamic router port and starts an aging timer for the port.

When receiving a membership report

A host sends an MLD report to the MLD querier in the following circumstances:

- If the host has been a member of an IPv6 multicast group, after receiving an MLD query, the host responds to the query with an MLD report.
- When the host wants to join an IPv6 multicast group, it sends an MLD report to the MLD querier, specifying the IPv6 multicast group to join.

After receiving an MLD report, the switch forwards it through all the router ports in the VLAN, resolves the address of the reported IPv6 multicast group, and performs the following judgment:

- If no forwarding entry matches the group address, the switch creates a forwarding entry for the group, adds the port that received the MLD report as a dynamic member port to the forwarding entry for the group, and starts an aging timer for the port.
- If a forwarding entry matches the group address, but the port that received the MLD report is not in the forwarding entry for the group, the switch adds the port as a dynamic member port to the forwarding entry, and starts an aging timer for the port.
- If a forwarding entry matches the group address and the port that received the MLD report is in the forwarding entry for the group, the switch restarts the aging timer for the port.

A switch does not forward an MLD report through a non-router port. The reason is that if the switch forwards a report message through a member port, all the attached hosts that monitor the reported IPv6

multicast address suppress their own reports after receiving this report according to the MLD report suppression mechanism. This prevents the switch from confirming whether the reported multicast group still has active members attached to that port.

For more information about the MLD report suppression mechanism of hosts, see "[Configuring MLD \(available only on the HP 5500 EI\)](#)."

When receiving a done message

When a host leaves an IPv6 multicast group, the host sends an MLD done message to the multicast routers. When the switch receives the MLD done message on a dynamic member port, the switch first checks whether a forwarding entry matches the IPv6 multicast group address in the message, and, if a match is found, whether the forwarding entry contains the dynamic member port.

- If no forwarding entry matches the IPv6 multicast group address, or if the forwarding entry does not contain the port, the switch directly discards the MLD done message.
- If a forwarding entry matches the IPv6 multicast group address and contains the port, the switch forwards the done message to all router ports in the native VLAN. Because the switch does not know whether any other hosts attached to the port are still listening to that IPv6 multicast group address, the switch does not immediately remove the port from the forwarding entry for that group. Instead, it restarts the aging timer for the port.

After receiving the MLD done message, the MLD querier resolves the IPv6 multicast group address in the message and sends an MLD multicast-address-specific query to that IPv6 multicast group through the port that received the MLD done message. After receiving the MLD multicast-address-specific query, the switch forwards it through all its router ports in the VLAN and all member ports of the IPv6 multicast group. The switch also performs the following judgment for the port that received the MLD done message:

- If the port (assuming that it is a dynamic member port) receives an MLD report in response to the MLD multicast-address-specific query before its aging timer expires, it indicates that some host attached to the port is receiving or expecting to receive IPv6 multicast data for that IPv6 multicast group. The switch restarts the aging timer for the port.
- If the port receives no MLD report in response to the MLD multicast-address-specific query before its aging timer expires, it indicates that no hosts attached to the port are still monitoring that IPv6 multicast group address. The switch removes the port from the forwarding entry for the IPv6 multicast group when the aging timer expires.

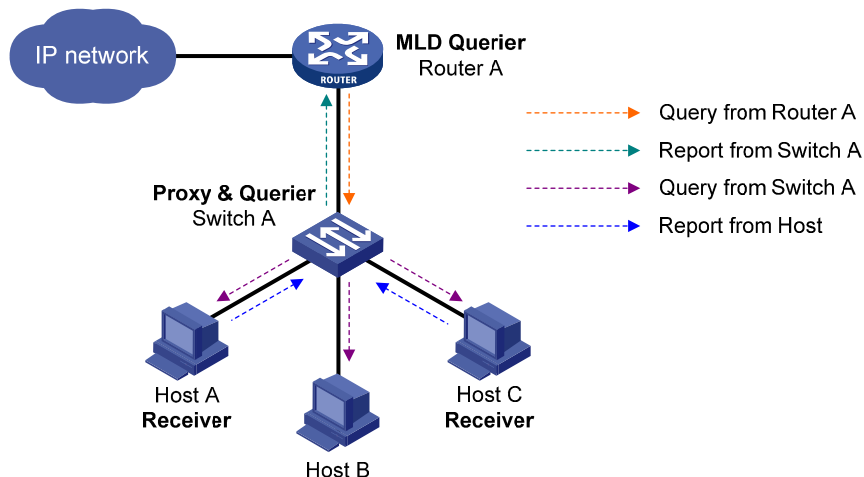
MLD snooping proxying

You can configure the MLD snooping proxying function on an edge device to reduce the number of MLD reports and done messages sent to its upstream device. The device configured with MLD snooping proxying is called an MLD snooping proxy. It is a host from the perspective of its upstream device.

NOTE:

Even though an MLD snooping proxy is a host from the perspective of its upstream device, the MLD membership report suppression mechanism for hosts does not take effect on it. For more information about the MLD report suppression mechanism for hosts, see "[Configuring MLD \(available only on the HP 5500 EI\)](#)."

Figure 67 Network diagram



As shown in [Figure 67](#), Switch A works as an MLD snooping proxy. As a host from the perspective of the querier Router A, Switch A represents its attached hosts to send their membership reports and done messages to Router A.

[Table 8](#) describes how an MLD snooping proxy processes MLD messages.

Table 8 MLD message processing on an MLD snooping proxy

MLD message	Actions
General query	When receiving an MLD general query, the proxy forwards it to all ports but the receiving port. In addition, the proxy generates a report according to the group memberships that it maintains and sends the report out of all router ports.
Multicast-address-specific query	In response to the MLD group-specific query for a certain IPv6 multicast group, the proxy sends the report to the group out of all router ports if the forwarding entry for the group still contains a member port.
Report	<p>When receiving a report for an IPv6 multicast group, the proxy looks up the multicast forwarding table for the entry for the multicast group.</p> <ul style="list-style-type: none"> • If a forwarding entry matches the IPv6 multicast group, and contains the receiving port as a dynamic member port, the proxy restarts the aging timer for the port. • If a forwarding entry matches the IPv6 multicast group but does not contain the receiving port, the proxy adds the port to the forwarding entry as a dynamic member port and starts an aging timer for the port. • If no forwarding entry matches the IPv6 multicast group, the proxy creates a forwarding entry for the group, adds the receiving port to the forwarding entry as a dynamic member port, and starts an aging timer for the port. <p>Then, the switch sends the report to the group out of all router ports.</p>
Done	In response to a done message for an IPv6 multicast group, the proxy sends a multicast-address-specific query for the group out of the receiving port. After making sure that no member port is contained in the forwarding entry for the IPv6 multicast group, the proxy sends a done message for the group out of all router ports.

Protocols and standards

RFC 4541, *Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches*

MLD snooping configuration task list

Task	Remarks	
Configuring basic MLD snooping functions	Enabling MLD snooping	Required
	Specifying the version of MLD snooping	Optional
	Configuring IPv6 static multicast MAC address entries	Optional
Configuring MLD snooping port functions	Configuring aging timers for dynamic ports	Optional
	Configuring static ports	Optional
	Configuring a port as a simulated member host	Optional
	Enabling fast-leave processing	Optional
	Disabling a port from becoming a dynamic router port	Optional
Configuring MLD snooping querier	Enabling MLD snooping querier	Optional
	Configuring parameters for MLD queries and responses	Optional
	Configuring the source IPv6 addresses for MLD queries	Optional
Configuring MLD snooping proxying	Enabling MLD snooping proxying	Optional
	Configuring the source IPv6 addresses for the MLD messages sent by the proxy	Optional
Configuring an MLD snooping policy	Configuring an IPv6 multicast group filter	Optional
	Configuring IPv6 multicast source port filtering	Optional
	Enabling dropping unknown IPv6 multicast data	Optional
	Configuring MLD report suppression	Optional
	Setting the maximum number of multicast groups that a port can join	Optional
	Enabling IPv6 multicast group replacement	Optional
	Setting the 802.1p precedence for MLD messages	Optional
	Configuring an IPv6 multicast user control policy	Optional
	Enabling the MLD snooping host tracking function	Optional
Setting the DSCP value for MLD messages	Optional	

For the configuration tasks in this section:

- In MLD-snooping view, configurations that you make are effective in all VLANs . In VLAN view, configurations that you make are effective only on the ports that belong to the current VLAN. For a given VLAN, a configuration that you make in MLD-snooping view is effective only if you do not make the same configuration in VLAN view.

- In MLD-snooping view, configurations that you make are effective on all ports. In Layer 2 Ethernet interface view or Layer 2 aggregate interface view, configurations that you make are effective only on the current port. In port group view, configurations that you make are effective on all ports in only the current port group. For a given port, a configuration that you make in MLD-snooping view is effective only if you do not make the same configuration in Layer 2 Ethernet interface view, Layer 2 aggregate interface view, or port group view.
- For MLD snooping, configurations that you make on a Layer 2 aggregate interface do not interfere with those made on its member ports, nor do they participate in aggregation calculations. Configurations that you make on a member port of the aggregate group will not take effect until the port leaves the aggregate group.

Configuring basic MLD snooping functions

Configuration prerequisites

Before you configure basic MLD snooping functions, complete the following tasks:

- Enable IPv6 forwarding.
- Configure the corresponding VLANs.
- Determine the version of MLD snooping.

Enabling MLD snooping

Configuration guidelines

- You must enable MLD snooping globally before you enable it for a VLAN.
- After you enable MLD snooping for a VLAN, you cannot enable MLD or IPv6 PIM on the corresponding VLAN interface, and vice versa.
- MLD snooping for a VLAN works only on the ports in this VLAN.

Configuration procedure

To enable MLD snooping:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable MLD snooping globally and enter MLD-snooping view.	mld-snooping	Disabled by default
3. Return to system view.	quit	N/A
4. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
5. Enable MLD snooping for the VLAN.	mld-snooping enable	Disabled by default

Specifying the version of MLD snooping

Configuration guidelines

Different versions of MLD snooping can process different versions of MLD messages:

- MLDv1 snooping can process MLDv1 messages, but flood MLDv2 messages in the VLAN instead of processing them.
- MLDv2 snooping can process MLDv1 and MLDv2 messages.

If you change MLDv2 snooping to MLDv1 snooping, the system clears all MLD snooping forwarding entries that are dynamically created, and also does the following:

- Keeps static MLDv2 snooping forwarding entries (*, G).
- Clears static MLDv2 snooping forwarding entries (S, G), which will be restored when MLDv1 snooping is changed back to MLDv2 snooping.

For more information about static joining, see "[Configuring static ports.](#)"

Configuration procedure

To specify the version of MLD snooping:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
3. Specify the version of MLD snooping.	mld-snooping version <i>version-number</i>	Version 1 by default

Configuring IPv6 static multicast MAC address entries

In Layer-2 multicast, a Layer-2 IPv6 multicast protocol (such as, MLD snooping) can dynamically add IPv6 multicast MAC address entries. Or, you can manually configure IPv6 multicast MAC address entries.

Configuration guidelines

The configuration that you make in system view is effective on the specified interfaces. The configuration that you make in interface view or port group view is effective only on the current interface or interfaces in the current port group.

Any legal IPv6 multicast MAC address except 3333-xxxx-xxxx (where x represents a hexadecimal number from 0 to F) can be manually added to the MAC address table. IPv6 multicast MAC addresses are the MAC addresses whose the least significant bit of the most significant octet is 1.

Configuration procedure

To configure an IPv6 static multicast MAC address entry in system view:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure a static multicast MAC address entry.	mac-address multicast <i>mac-address interface interface-list</i> vlan <i>vlan-id</i>	No static multicast MAC address entries exist by default.

To configure an IPv6 static multicast MAC address entry in interface view:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet interface view, Layer 2 aggregate interface view, or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	In Ethernet interface view or Layer 2 aggregate interface view, the configuration is effective on only the current interface. In port group view, the configuration is effective on all ports in the port group.
3. Configure a static multicast MAC address entry.	mac-address multicast <i>mac-address</i> vlan <i>vlan-id</i>	No static multicast MAC address entries exist by default.

For more information about the **mac-address multicast** command, see *IP Multicast Command Reference*.

Configuring MLD snooping port functions

Configuration prerequisites

Before you configure MLD snooping port functions, complete the following tasks:

- Enable MLD snooping in the VLAN.
- Configure the corresponding port groups.
- Determine the aging time of dynamic router ports.
- Determine the aging time of dynamic member ports.
- Determine the IPv6 multicast group and IPv6 multicast source addresses.

Configuring aging timers for dynamic ports

If a switch receives no MLD general queries or IPv6 PIM hello messages on a dynamic router port when the aging timer of the port expires, the switch removes the port from the router port list.

If the switch receives no MLD reports for an IPv6 multicast group on a dynamic member port when the aging timer of the port expires, the switch removes the port from the forwarding entry for the IPv6 multicast group.

If the memberships of IPv6 multicast groups change frequently, you can set a relatively small value for the aging timer of the dynamic member ports. If the memberships of IPv6 multicast groups change rarely, you can set a relatively large value.

Setting the global aging timers for dynamic ports

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter MLD-snooping view.	mld-snooping	N/A
3. Set the global aging timer for dynamic router ports.	router-aging-time <i>interval</i>	260 seconds by default

Step	Command	Remarks
4.	Set the global aging timer for dynamic member ports. host-aging-time <i>interval</i>	260 seconds by default

Setting the aging timers for the dynamic ports in a VLAN

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Enter VLAN view. vlan <i>vlan-id</i>	N/A
3.	Set the aging timer for the dynamic router ports. mld-snooping router-aging-time <i>interval</i>	260 seconds by default
4.	Set the aging timer for the dynamic member ports. mld-snooping host-aging-time <i>interval</i>	260 seconds by default

Configuring static ports

Configuration guidelines

If all hosts attached to a port are interested in the IPv6 multicast data addressed to a particular IPv6 multicast group, configure the port as a static member port for that IPv6 multicast group.

You can configure a port as a static router port, through which the switch can forward all IPv6 multicast data that it received.

A static member port does not respond to queries from the MLD querier; when you configure a port as a static member port or cancel this configuration on the port, the port does not send an unsolicited MLD report or an MLD done message.

Static member ports and static router ports never age out. To remove such a port, you use the corresponding **undo** command.

Configuration procedure

To configure static ports:

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Enter Layer 2 Ethernet interface view, Layer 2 aggregate interface view, or port group view. <ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3.	Configure the port as a static member port. mld-snooping static-group <i>ipv6-group-address</i> [source-ip <i>ipv6-source-address</i>] vlan <i>vlan-id</i>	No static member ports exist by default.

Step	Command	Remarks
4. Configure the port as a static router port.	mld-snooping static-router-port vlan <i>vlan-id</i>	No static router ports exist by default.

Configuring a port as a simulated member host

Generally, a host that runs MLD can respond to MLD queries. If a host fails to respond, the multicast router might deem that the IPv6 multicast group has no members on the subnet, and removes the corresponding forwarding path.

To avoid this situation, you can configure a port on the switch as a simulated member host for an IPv6 multicast group. A simulated host is equivalent to an independent host. For example, when a simulated member host receives an MLD query, it gives a response separately. Therefore, the switch can continue receiving IPv6 multicast data.

A simulated host acts like a real host in the following ways:

- When a port is configured as a simulated member host, the switch sends an unsolicited MLD report through the port, and can respond to MLD general queries with MLD reports through the port.
- When the simulated joining configuration is canceled on the port, the switch sends an MLD done message through that port.

To configure a port as a simulated member host:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet interface view, Layer 2 aggregate interface view, port group view.	<ul style="list-style-type: none"> • Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> • Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Configure the port as a simulated member host.	mld-snooping host-join <i>ipv6-group-address</i> [source-ip <i>ipv6-source-address</i>] vlan <i>vlan-id</i>	Not configured by default.

NOTE:

Unlike a static member port, a port that you configure as a simulated member host ages out like a dynamic member port.

Enabling fast-leave processing

The fast-leave processing feature enables the switch to process MLD done messages quickly. After the fast-leave processing feature is enabled, when the switch receives an MLD done message on a port, it immediately removes that port from the forwarding entry for the multicast group specified in the message. Then, when the switch receives MLD multicast-address-specific queries for that multicast group, it does not forward them to that port.

On a port that has only one host attached, you can enable fast-leave processing to save bandwidth and resources. However, on a port that has multiple hosts attached, you should not enable fast-leave processing if you have enabled dropping unknown IPv6 multicast data globally or for the port. Otherwise, if a host on the port leaves an IPv6 multicast group, the other hosts attached to the port in the same IPv6 multicast group cannot receive the IPv6 multicast data for the group.

Enabling fast-leave processing globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter MLD-snooping view.	mld-snooping	N/A
3. Enable fast-leave processing.	fast-leave [vlan <i>vlan-list</i>]	Disabled by default

Enabling fast-leave processing on a port

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet interface view, Layer 2 aggregate interface view, or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Enable fast-leave processing.	mld-snooping fast-leave [vlan <i>vlan-list</i>]	Disabled by default.

Disabling a port from becoming a dynamic router port

The following problems exist in a multicast access network:

- After receiving an MLD general query or IPv6 PIM hello message from a connected host, a router port becomes a dynamic router port. Before its timer expires, this dynamic router port receives all multicast packets within the VLAN where the port belongs, and forwards them to the host, affecting normal multicast reception of the host.
- In addition, the MLD general query and IPv6 PIM hello message that the host sends affects the multicast routing protocol state on Layer 3 devices, such as the MLD querier or DR election, and might further cause network interruption.

To solve these problems, disable that router port from becoming a dynamic router port after the port receives an MLD general query or IPv6 PIM hello message, so as to improve network security and control over multicast users.

To disable a port from becoming a dynamic router port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter Layer 2 Ethernet interface view, Layer 2 aggregate interface view, or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Disable the port from becoming a dynamic router port.	mld-snooping router-port-deny [vlan <i>vlan-list</i>]	By default, a port can become a dynamic router port.

NOTE:

This configuration does not affect the static router port configuration.

Configuring MLD snooping querier

Configuration prerequisites

Before you configure MLD snooping querier, complete the following tasks:

- Enable MLD snooping in the VLAN.
- Determine the MLD general query interval.
- Determine the MLD last-member query interval.
- Determine the maximum response time for MLD general queries.
- Determine the source IPv6 address of MLD general queries.
- Determine the source IPv6 address of MLD multicast-address-specific queries.

Enabling MLD snooping querier

In an IPv6 multicast network that runs MLD, a multicast router or Layer 3 multicast switch sends MLD queries, so that all Layer 3 multicast devices can establish and maintain multicast forwarding entries, in order to forward multicast traffic correctly at the network layer. This router or Layer 3 switch is called the "MLD querier." For more information about MLD querier, see "[Configuring MLD \(available only on the HP 5500 E1\)](#)."

However, a Layer 2 multicast switch does not support MLD. Therefore, it cannot send MLD general queries by default. When you enable MLD snooping querier on a Layer 2 switch in a VLAN where multicast traffic is switched only at Layer 2 and no Layer 3 multicast devices are present, the Layer 2 switch sends MLD queries, so that multicast forwarding entries can be created and maintained at the data link layer.

! **IMPORTANT:**

It is meaningless to configure an MLD snooping querier in an IPv6 multicast network that runs MLD. Although an MLD snooping querier does not participate in MLD querier elections, it might affect MLD querier elections because it sends MLD general queries with a low source IPv6 address.

To enable the MLD snooping querier:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
3. Enable the MLD snooping querier.	mld-snooping querier	Disabled by default

Configuring parameters for MLD queries and responses

Configuration guidelines

You can modify the MLD general query interval based on the actual condition of the network.

A multicast listening host starts a timer for each IPv6 multicast group that it has joined when it receives an MLD query (general query or multicast-address-specific query). This timer is initialized to a random value in the range of 0 to the maximum response delay advertised in the MLD query message. When the timer value decreases to 0, the host sends an MLD report to the IPv6 multicast group.

To speed up the response of hosts to MLD queries and avoid simultaneous timer expirations causing MLD report traffic bursts, you must properly set the maximum response delay.

- The maximum response delay for MLD general queries is set by the **max-response-time** command.
- The maximum response delay for MLD multicast-address-specific queries equals the MLD last-listener query interval.

In the configuration, make sure that the interval for sending MLD general queries is greater than the maximum response delay for MLD general queries. Otherwise, undesired deletion of IPv6 multicast members might occur.

Configuration procedure

To configure MLD queries and responses globally:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter MLD-snooping view.	mld-snooping	N/A
3. Set the maximum response delay for MLD general queries.	max-response-time <i>interval</i>	10 seconds by default
4. Set the MLD last-member query interval.	last-listener-query-interval <i>interval</i>	1 second by default

To configure the parameters for MLD queries and responses in a VLAN

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
3. Set the MLD query interval.	mld-snooping query-interval <i>interval</i>	125 seconds by default

Step	Command	Remarks
4. Set the maximum response delay for MLD general queries.	mld-snooping max-response-time <i>interval</i>	10 seconds by default
5. Set the MLD last-member query interval.	mld-snooping last-listener-query-interval <i>interval</i>	1 second by default

Configuring the source IPv6 addresses for MLD queries

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
3. Configure the source IPv6 address of MLD general queries.	mld-snooping general-query source-ip { <i>ipv6-address</i> current-interface }	FE80::02FF:FFFF:FE00:0001 by default
4. Configure the source IPv6 address of MLD multicast-address-specific queries.	mld-snooping special-query source-ip { <i>ipv6-address</i> current-interface }	FE80::02FF:FFFF:FE00:0001 by default



IMPORTANT:

The source IPv6 address of MLD query messages might affect MLD querier election within the subnet.

Configuring MLD snooping proxying

Configuration prerequisites

Before you configure MLD snooping proxying in a VLAN, complete the following tasks:

- Enable MLD snooping in the VLAN.
- Determine the source IPv6 address for the MLD reports sent by the proxy.
- Determine the source IPv6 address for the MLD done messages sent by the proxy.

Enabling MLD snooping proxying

The MLD snooping proxying function works on a per-VLAN basis. After you enable the function in a VLAN, the device works as the MLD snooping proxy for the downstream hosts and upstream router in the VLAN.

To enable MLD snooping proxying in a VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A

Step	Command	Remarks
3.	Enable MLD snooping proxying in the VLAN. mld-snooping proxying enable	Disabled by default

Configuring the source IPv6 addresses for the MLD messages sent by the proxy

You can set the source IPv6 addresses for the MLD reports and done messages that the MLD snooping proxy sends on behalf of its attached hosts.

To configure the source IPv6 addresses for the MLD messages that the MLD snooping proxy sends in a VLAN:

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Enter VLAN view. vlan <i>vlan-id</i>	N/A
3.	Configure a source IPv6 address for the MLD reports that the proxy sends. mld-snooping report source-ip { <i>ipv6-address</i> current-interface }	The default is FE80::02FF:FFFF:FE00:0001.
4.	Configure a source IPv6 address for the MLD done messages that the proxy sends. mld-snooping done source-ip { <i>ipv6-address</i> current-interface }	The default is FE80::02FF:FFFF:FE00:0001.

Configuring an MLD snooping policy

Configuration prerequisites

Before you configure an MLD snooping policy, complete the following tasks:

- Enable MLD snooping in the VLAN.
- Determine the IPv6 ACL rule for IPv6 multicast group filtering.
- Determine the maximum number of IPv6 multicast groups that a port can join.
- Determine the 802.1p precedence for MLD messages.

Configuring an IPv6 multicast group filter

On an MLD snooping-enabled switch, you can configure an IPv6 multicast group filter to limit multicast programs available to users.

Configuration guidelines

In an application, when a user requests a multicast program, the user's host initiates an MLD report. After receiving this report message, the switch resolves the IPv6 multicast group address in the report and looks up the ACL. If a match is found to permit the port that received the report to join the IPv6 multicast group, the switch creates an MLD snooping forwarding entry for the IPv6 multicast group and adds the port to

the forwarding entry. Otherwise, the switch drops this report message, in which case, the IPv6 multicast data for the IPv6 multicast group is not sent to this port, and the user cannot retrieve the program.

When you configure a multicast group filter in an IPv6 multicast VLAN, be sure to configure the filter in the sub-VLANs of the IPv6 multicast VLAN. Otherwise, the configuration does not take effect.

In MLDv2, when a host is enabled to join multiple multicast groups, the multicast group filter cannot correctly filter multicast groups because the host that runs MLDv2 sends multiple multicast groups that it wants to join in one membership report.

Configuration procedure

To configure an IPv6 multicast group globally:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter MLD-snooping view.	mld-snooping	N/A
3. Configure an IPv6 multicast group filter.	group-policy <i>acl6-number</i> [vlan <i>vlan-list</i>]	By default, no IPv6 group filter is globally configured. That is, the hosts in a VLAN can join any valid multicast group.

To configure an IPv6 multicast group filter for a port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet interface view, Layer 2 aggregate interface view, or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Configure an IPv6 multicast group filter.	mld-snooping group-policy <i>acl6-number</i> [vlan <i>vlan-list</i>]	By default, no IPv6 group filter is configured on an interface. That is, the hosts on the interface can join any valid multicast group.

Configuring IPv6 multicast source port filtering

When the IPv6 multicast source port filtering feature is enabled on a port, the port can connect only to IPv6 multicast receivers rather than multicast sources. The reason is that the port blocks all IPv6 multicast data packets but it permits multicast protocol packets to pass.

If this feature is disabled on a port, the port can connect to both multicast sources and IPv6 multicast receivers.

Configuring IPv6 multicast source port filtering globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter MLD-snooping view.	mld-snooping	N/A
3. Enable IPv6 multicast source port filtering.	source-deny port <i>interface-list</i>	Disabled by default

Configuring IPv6 multicast source port filtering for a port

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet interface view, or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view: interface <i>interface-type interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Enable IPv6 multicast source port filtering.	mld-snooping source-deny	Disabled by default.

NOTE:

Some models of devices, when enabled to filter IPv6 multicast data based on the source ports, are automatically enabled to filter IPv4 multicast data based on the source ports.

Enabling dropping unknown IPv6 multicast data

Configuration guidelines

Unknown IPv6 multicast data refers to IPv6 multicast data for which no entries exist in the MLD snooping forwarding table. When the switch receives such IPv6 multicast traffic, one of the following occurs:

- When the function of dropping unknown IPv6 multicast data is disabled, the switch floods unknown IPv6 multicast data in the VLAN to which the unknown IPv6 multicast data belongs.
- When the function of dropping unknown IPv6 multicast data is enabled, the switch forwards unknown multicast data to its router ports instead of flooding it in the VLAN. If no router ports exist, the switch drops the unknown multicast data.

Configuration procedure

To enable dropping unknown IPv6 multicast data in a VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
3. Enable dropping unknown IPv6 multicast data.	mld-snooping drop-unknown	Disabled by default

Configuring MLD report suppression

Configuration guidelines

When a Layer 2 switch receives an MLD report from an IPv6 multicast group member, the Layer 2 switch forwards the message to the Layer 3 device that directly connects to the Layer 2 switch. When multiple members of an IPv6 multicast group are attached to the Layer 2 switch, the Layer 3 device might receive duplicate MLD reports for the IPv6 multicast group from these members.

With the MLD report suppression function enabled, within a query interval, the Layer 2 switch forwards only the first MLD report for the IPv6 multicast group to the Layer 3 device. It does not forward subsequent MLD reports for the same IPv6 multicast group to the Layer 3 device. This helps reduce the number of packets being transmitted over the network.

On an MLD snooping proxy, MLD reports for an IPv6 multicast group from downstream hosts are suppressed if the forwarding entry for the multicast group exists on the proxy, whether the suppression function is enabled or not.

Configuration procedure

To configure MLD report suppression:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter MLD-snooping view.	mld-snooping	N/A
3. Enable MLD report suppression.	report-aggregation	Enabled by default

Setting the maximum number of multicast groups that a port can join

You can set the maximum number of IPv6 multicast groups that a port can join to regulate the traffic on the port.

When you configure this maximum number, if the number of IPv6 multicast groups the port has joined exceeds the configured maximum value, the system deletes all the forwarding entries for the port from the MLD snooping forwarding table, and the hosts on this port join IPv6 multicast groups again until the number of IPv6 multicast groups that the port joins reaches the maximum value. When the port joins an IPv6 multicast group, if the port has been configured as a static member port, the system applies the configurations to the port again. If you have configured simulated joining on the port, the system establishes corresponding forwarding entry for the port after receiving a report from the simulated member host.

To configure the maximum number of IPv6 multicast groups that a port can join:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter Layer 2 Ethernet interface view, Layer 2 aggregate interface view, port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Set the maximum number of IPv6 multicast groups that a port can join.	mld-snooping group-limit <i>limit</i> [vlan <i>vlan-list</i>]	1000 by default.

Enabling IPv6 multicast group replacement

For various reasons, the number of IPv6 multicast groups that a switch or a port can join might exceed the upper limit. In addition, in some specific applications, an IPv6 multicast group that the switch newly joins must replace an existing IPv6 multicast group automatically. A typical example is channel switching. To view a new TV channel, a user switches from the current IPv6 multicast group to the new one.

To realize such requirements, you can enable the IPv6 multicast group replacement function on the switch or on a certain port. When the number of IPv6 multicast groups that the switch or the port has joined reaches the limit, one of the following occurs:

- If the IPv6 multicast group replacement feature is disabled, new MLD reports are automatically discarded.
- If the IPv6 multicast group replacement feature is enabled, the IPv6 multicast group that the switch or the port newly joins automatically replaces an existing IPv6 multicast group that has the lowest IPv6 address.

⚠ IMPORTANT:

Be sure to configure the maximum number of IPv6 multicast groups allowed on a port (see "[Setting the maximum number of multicast groups that a port can join](#)") before enabling IPv6 multicast group replacement. Otherwise, the IPv6 multicast group replacement functionality will not take effect.

Enabling IPv6 multicast group replacement globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter MLD-snooping view.	mld-snooping	N/A
3. Enable IPv6 multicast group replacement.	overflow-replace [vlan <i>vlan-list</i>]	Disabled by default

Enabling IPv6 multicast group replacement for a port

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter Layer 2 Ethernet interface view, Layer 2 aggregate interface view, or port group view.	<ul style="list-style-type: none"> Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view: interface <i>interface-type interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Enable IPv6 multicast group replacement.	mld-snooping overflow-replace [vlan <i>vlan-list</i>]	Disabled by default.

Setting the 802.1p precedence for MLD messages

You can change the 802.1p precedence of MLD messages so that they can be assigned higher forwarding priority when congestion occurs on their outgoing ports.

Setting the 802.1p precedence for MLD messages globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter MLD-snooping view.	mld-snooping	N/A
3. Set the 802.1p precedence for MLD messages.	dot1p-priority <i>priority-number</i>	The default 802.1p precedence for MLD messages is 0.

Setting the 802.1p precedence for MLD messages in a VLAN

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
3. Set the 802.1p precedence for MLD messages.	mld-snooping dot1p-priority <i>priority-number</i>	The default 802.1p precedence for MLD messages is 0.

Configuring an IPv6 multicast user control policy

IPv6 multicast user control policies are configured on access switches to allow only authorized users to receive requested IPv6 multicast data. This helps restrict users from ordering certain multicast-on-demand programs.

Configuration guidelines

In practice, a device first needs to perform authentication (for example, 802.1X authentication) for the connected hosts through a RADIUS server. Then, the device uses the configured multicast user control policy to perform multicast access control for authenticated users as follows.

- After receiving an MLD report from a host, the access switch matches the IPv6 multicast group address and multicast source address carried in the report with the configured policies. If a match

is found, the user is allowed to join the multicast group. Otherwise, the join report is dropped by the access switch.

- After receiving a done message from a host, the access switch matches the IPv6 multicast group address and source address against the policies. If a match is found, the host is allowed to leave the group. Otherwise, the done message is dropped by the access switch.

An IPv6 multicast user control policy is functionally similar to an IPv6 multicast group filter. A difference lies in that a control policy can control both multicast joining and leaving of users based on authentication and authorization, but a multicast group filter is configured on a port to control only multicast joining but not leaving of users without authentication or authorization.

Configuration procedure

To configure a multicast user control policy

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a user profile and enter its view.	user-profile <i>profile-name</i>	N/A
3. Configure a multicast user control policy.	mld-snooping access-policy <i>acl6-number</i>	No policy is configured by default. That is, a host can join or leave a valid multicast group at any time.
4. Return to system view.	quit	N/A
5. Enable the created user profile.	user-profile <i>profile-name</i> enable	Not enabled by default.

For more information about the **user-profile** and **user-profile enable** commands, see *Security Command Reference*.

Enabling the MLD snooping host tracking function

With the MLD snooping host tracking function, the switch can record the information of the member hosts that are receiving IPv6 multicast traffic, including the host IPv6 address, running duration, and timeout time. You can monitor and manage the member hosts according to the recorded information.

Enabling the MLD snooping host tracking function globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter MLD-snooping view.	mld-snooping	N/A
3. Enable the MLD snooping host tracking function globally.	host-tracking	Disabled by default

Enabling the MLD snooping host tracking function in a VLAN

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter VLAN view.	vlan <i>vlan-id</i>	N/A

Step	Command	Remarks
3.	Enable the MLD snooping host tracking function in the VLAN. mld-snooping host-tracking	Disabled by default

Setting the DSCP value for MLD messages

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Enter MLD-snooping view. mld-snooping	N/A
3.	Set the DSCP value for MLD messages. dscp dscp-value	By default, the DSCP value in MLD messages is 48.

NOTE:

This configuration applies to only the MLD messages that the local switch generates rather than those forwarded ones.

Displaying and maintaining MLD snooping

Task	Command	Remarks
Display MLD snooping group information.	display mld-snooping group [vlan <i>vlan-id</i>] [slot <i>slot-number</i>] [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about the hosts tracked by MLD snooping.	display mld-snooping host vlan <i>vlan-id</i> group <i>ipv6-group-address</i> [source <i>ipv6-source-address</i>] [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 static multicast MAC address entries.	display mac-address [<i>mac-address</i> [vlan <i>vlan-id</i>] [multicast] [vlan <i>vlan-id</i>] [count]] [{ begin exclude include } <i>regular-expression</i>]	Available in user view
Display statistics for the MLD messages learned through MLD snooping.	display mld-snooping statistics [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Remove dynamic group entries of a specified MLD snooping group or all MLD snooping groups.	reset mld-snooping group { <i>ipv6-group-address</i> all } [vlan <i>vlan-id</i>]	Available in user view
Clear statistics for the MLD messages learned through MLD snooping.	reset mld-snooping statistics	Available in user view

NOTE:

- The **reset mld-snooping group** command works only on an MLD snooping-enabled VLAN, but not in a VLAN with MLD enabled on its VLAN interface.
 - The **reset mld-snooping group** command cannot remove the static group entries of MLD snooping groups.
-

For more information about the **display mac-address multicast** command, see *IP Multicast Command Reference*.

MLD snooping configuration examples

IPv6 group policy and simulated joining configuration example

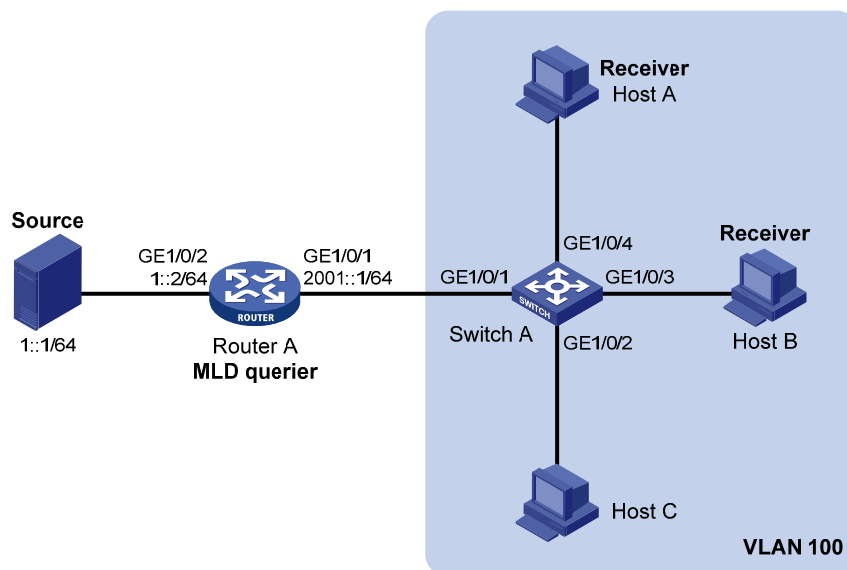
Network requirements

As shown in [Figure 68](#), MLDv1 runs on Router A, MLDv1 snooping required on Switch A, and Router A acts as the MLD querier on the subnet.

The receivers, Host A and Host B can receive IPv6 multicast traffic addressed to IPv6 multicast group FF1E::101 only.

IPv6 multicast data for group FF1E::101 can be forwarded through GigabitEthernet 1/0/3 and GigabitEthernet 1/0/4 of Switch A even if Host A and Host B accidentally, temporarily stop receiving IPv6 multicast data, and that Switch A drops unknown IPv6 multicast data and does not broadcast the data to the VLAN where Switch A resides.

Figure 68 Network diagram



Configuration procedure

1. Enable IPv6 forwarding and configure IPv6 addresses:
Enable IPv6 forwarding and configure an IPv6 address and prefix length for each interface as per [Figure 68](#). (Details not shown.)
2. Configure Router A:

Enable IPv6 multicast routing, enable IPv6 PIM-DM on each interface, and enable MLDv1 on GigabitEthernet 1/0/1.

```
<RouterA> system-view
[RouterA] multicast ipv6 routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] mld enable
[RouterA-GigabitEthernet1/0/1] pim ipv6 dm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] pim ipv6 dm
[RouterA-GigabitEthernet1/0/2] quit
```

3. Configure Switch A:

Enable MLD snooping globally.

```
<SwitchA> system-view
[SwitchA] mld-snooping
[SwitchA-mld-snooping] quit
```

Create VLAN 100, assign GigabitEthernet 1/0/1 through GigabitEthernet 1/0/4 to this VLAN, and enable MLD snooping and the function of dropping IPv6 unknown multicast traffic in the VLAN.

```
[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/4
[SwitchA-vlan100] mld-snooping enable
[SwitchA-vlan100] mld-snooping drop-unknown
[SwitchA-vlan100] quit
```

Configure an IPv6 multicast group filter so that the hosts in VLAN 100 can join only the IPv6 multicast group FF1E::101.

```
[SwitchA] acl ipv6 number 2001
[SwitchA-acl6-basic-2001] rule permit source ff1e::101 128
[SwitchA-acl6-basic-2001] quit
[SwitchA] mld-snooping
[SwitchA-mld-snooping] group-policy 2001 vlan 100
[SwitchA-mld-snooping] quit
```

Configure GigabitEthernet 1/0/3 and GigabitEthernet 1/0/4 as simulated hosts for IPv6 multicast group FF1E::101.

```
[SwitchA] interface gigabitethernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] mld-snooping host-join ff1e::101 vlan 100
[SwitchA-GigabitEthernet1/0/3] quit
[SwitchA] interface gigabitethernet 1/0/4
[SwitchA-GigabitEthernet1/0/4] mld-snooping host-join ff1e::101 vlan 100
[SwitchA-GigabitEthernet1/0/4] quit
```

4. Verify the configuration:

Display detailed MLD snooping group information in VLAN 100 on Switch A.

```
[SwitchA] display mld-snooping group vlan 100 verbose
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
```

```

Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port
Subvlan flags: R-Real VLAN, C-Copy VLAN
Vlan(id):100.
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
  Router port(s):total 1 port(s).
    GE1/0/1                (D) ( 00:01:30 )
IP group(s):the following ip group(s) match to one mac group.
  IP group address:FF1E::101
  (::, FF1E::101):
  Attribute:      Host Port
  Host port(s):total 2 port(s).
    GE1/0/3                (D) ( 00:03:23 )
    GE1/0/4                (D) ( 00:04:10 )
MAC group(s):
  MAC group address:3333-0000-0101
  Host port(s):total 2 port(s).
    GE1/0/3
    GE1/0/4

```

The output shows that GigabitEthernet 1/0/3 and GigabitEthernet 1/0/4 of Switch A have joined IPv6 multicast group FF1E::101.

Static port configuration example

Network requirements

As shown in [Figure 69](#), MLDv1 runs on Router A, and MLDv1 snooping runs on Switch A, Switch B and Switch C. Router A acts as the MLD querier.

Host A and Host C are permanent receivers of IPv6 multicast group FF1E::101. GigabitEthernet 1/0/3 and GigabitEthernet 1/0/5 on Switch C are required to be configured as static member ports for multicast group FF1E::101 to enhance the reliability of multicast traffic transmission.

Suppose STP runs on the network. To avoid data loops, the forwarding path from Switch A to Switch C is blocked under normal conditions, and IPv6 multicast traffic flows to the receivers attached to Switch C only along the path of Switch A—Switch B—Switch C.

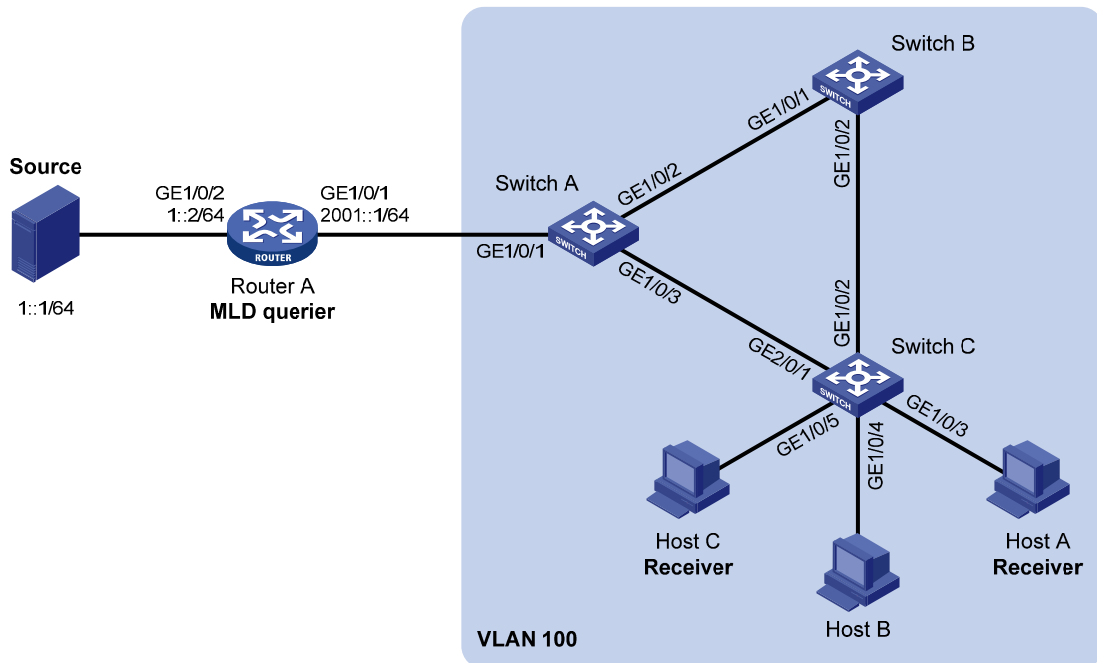
Configure GigabitEthernet 1/0/3 on Switch C as a static router port, so that IPv6 multicast traffic can flow to the receivers nearly uninterruptedly along the path of Switch A—Switch C in the case that the path of Switch A—Switch B—Switch C becomes blocked.

NOTE:

If no static router port is configured, when the path of Switch A—Switch B—Switch C becomes blocked, at least one MLD query-response cycle must be completed before the IPv6 multicast data can flow to the receivers along the new path of Switch A—Switch C. Namely, IPv6 multicast delivery will be interrupted during this process.

For more information about the Spanning Tree Protocol (STP), see *Layer 2—LAN Switching Configuration Guide*.

Figure 69 Network diagram



Configuration procedure

1. Enable IPv6 forwarding and configure IPv6 addresses:
Enable IPv6 forwarding and configure an IPv6 address and prefix length for each interface as per Figure 69.
2. Configure Router A:
Enable IPv6 multicast routing, enable IPv6 PIM-DM on each interface, and enable MLD on GigabitEthernet 1/0/1.

```

<RouterA> system-view
[RouterA] multicast ipv6 routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] mld enable
[RouterA-GigabitEthernet1/0/1] pim ipv6 dm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] pim ipv6 dm
[RouterA-GigabitEthernet1/0/2] quit

```
3. Configure Switch A:
Enable MLD snooping globally.

```

<SwitchA> system-view
[SwitchA] mld-snooping
[SwitchA-mld-snooping] quit

```

Create VLAN 100, assign GigabitEthernet 1/0/1 through GigabitEthernet 1/0/3 to this VLAN, and enable MLD snooping in the VLAN.

```

[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/3
[SwitchA-vlan100] mld-snooping enable

```

```
[SwitchA-vlan100] quit
# Configure GigabitEthernet 1/0/3 to be a static router port.
[SwitchA] interface gigabitEthernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] mld-snooping static-router-port vlan 100
[SwitchA-GigabitEthernet1/0/3] quit
```

4. Configure Switch B:

Enable MLD snooping globally.

```
<SwitchB> system-view
[SwitchB] mld-snooping
[SwitchB-mld-snooping] quit
```

Create VLAN 100, assign GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 to this VLAN, and enable MLD snooping in the VLAN.

```
[SwitchB] vlan 100
[SwitchB-vlan100] port gigabitEthernet 1/0/1 gigabitEthernet 1/0/2
[SwitchB-vlan100] mld-snooping enable
[SwitchB-vlan100] quit
```

5. Configure Switch C:

Enable MLD snooping globally.

```
<SwitchC> system-view
[SwitchC] mld-snooping
[SwitchC-mld-snooping] quit
```

Create VLAN 100, assign GigabitEthernet 1/0/1 through GigabitEthernet 1/0/5 to this VLAN, and enable MLD snooping in the VLAN.

```
[SwitchC] vlan 100
[SwitchC-vlan100] port gigabitEthernet 1/0/1 to gigabitEthernet 1/0/5
[SwitchC-vlan100] mld-snooping enable
[SwitchC-vlan100] quit
```

Configure GigabitEthernet 1/0/3 and GigabitEthernet 1/0/5 as static member ports for IPv6 multicast group FF1E::101.

```
[SwitchC] interface GigabitEthernet 1/0/3
[SwitchC-GigabitEthernet1/0/3] mld-snooping static-group ff1e::101 vlan 100
[SwitchC-GigabitEthernet1/0/3] quit
[SwitchC] interface GigabitEthernet 1/0/5
[SwitchC-GigabitEthernet1/0/5] mld-snooping static-group ff1e::101 vlan 100
[SwitchC-GigabitEthernet1/0/5] quit
```

6. Verify the configuration:

Display detailed MLD snooping group information in VLAN 100 on Switch A.

```
[SwitchA] display mld-snooping group vlan 100 verbose
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
```

Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port

Subvlan flags: R-Real VLAN, C-Copy VLAN

Vlan(id):100.

```
  Total 1 IP Group(s).
```

```

Total 1 IP Source(s).
Total 1 MAC Group(s).
Router port(s):total 2 port(s).
    GE1/0/1          (D) ( 00:01:30 )
    GE1/0/3          (S)
IP group(s):the following ip group(s) match to one mac group.
IP group address:FF1E::101
  (::, FF1E::101):
  Attribute:      Host Port
  Host port(s):total 1 port(s).
    GE1/0/2          (D) ( 00:03:23 )
MAC group(s):
MAC group address:3333-0000-0101
Host port(s):total 1 port(s).
  GE1/0/2

```

The output shows that GigabitEthernet 1/0/3 of Switch A has become a static router port.

Display detailed MLD snooping group information in VLAN 100 on Switch C.

```
[SwitchC] display mld-snooping group vlan 100 verbose
```

```

Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).

Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port
Subvlan flags: R-Real VLAN, C-Copy VLAN
Vlan(id):100.
Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).
Router port(s):total 1 port(s).
    GE1/0/2          (D) ( 00:01:23 )
IP group(s):the following ip group(s) match to one mac group.
IP group address:FF1E::101
  (::, FF1E::101):
  Attribute:      Host Port
  Host port(s):total 2 port(s).
    GE1/0/3          (S)
    GE1/0/5          (S)
MAC group(s):
MAC group address:3333-0000-0101
Host port(s):total 2 port(s).
  GE1/0/3
  GE1/0/5

```

The output shows that GigabitEthernet 1/0/3 and GigabitEthernet 1/0/5 on Switch C have become static member ports for IPv6 multicast group FF1E::101.

MLD snooping querier configuration example

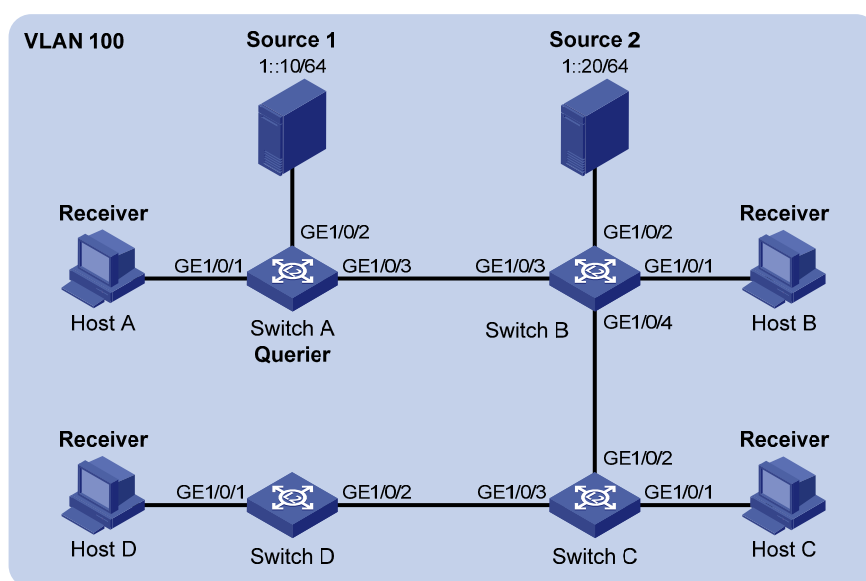
Network requirements

As shown in Figure 70, in a Layer-2-only network environment, two multicast sources (Source 1 and Source 2) send IPv6 multicast data to multicast groups FF1E::101 and FF1E::102 respectively, Host A and Host C are receivers of multicast group FF1E::101 and Host B and Host D are receivers of multicast group FF1E::102.

MLDv1 runs on all the receivers and MLDv1 snooping runs on all the switches. Switch A, which is close to the multicast sources, is chosen as the MLD snooping querier.

To prevent flooding of unknown multicast traffic within the VLAN, configure all the switches to drop unknown multicast data packets.

Figure 70 Network diagram



Configuration procedure

1. Configure Switch A:

Enable IPv6 forwarding and enable MLD snooping globally.

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] mld-snooping
[SwitchA-mld-snooping] quit
```

Create VLAN 100 and assign GigabitEthernet 1/0/1 through GigabitEthernet 1/0/3 to VLAN 100.

```
[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/3
```

Enable MLD snooping and the function of dropping unknown IPv6 multicast data packets in VLAN 100.

```
[SwitchA-vlan100] mld-snooping enable
[SwitchA-vlan100] mld-snooping drop-unknown
```

Configure MLD snooping querier feature in VLAN 100.

```
[SwitchA-vlan100] mld-snooping querier
```

```
[SwitchA-vlan100] quit
```

2. Configure Switch B:

Enable IPv6 forwarding and enable MLD snooping globally.

```
<SwitchB> system-view
```

```
[SwitchB] ipv6
```

```
[SwitchB] mld-snooping
```

```
[SwitchB-mld-snooping] quit
```

Create VLAN 100, add GigabitEthernet 1/0/1 through GigabitEthernet 1/0/4 into VLAN 100.

```
[SwitchB] vlan 100
```

```
[SwitchB-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/4
```

Enable the MLD snooping feature and the function of dropping unknown IPv6 multicast data packets in VLAN 100.

```
[SwitchB-vlan100] mld-snooping enable
```

```
[SwitchB-vlan100] mld-snooping drop-unknown
```

```
[SwitchB-vlan100] quit
```

Configurations of Switch C and Switch D are similar to the configuration of Switch B.

3. Verify the configuration:

When the MLD snooping querier starts to work, all the switches but the querier receive MLD general queries. Use the **display mld-snooping statistics** command to view the statistics information of these MLD messages received.

Display the MLD message statistics on Switch B.

```
[SwitchB-vlan100] display mld-snooping statistics
```

```
Received MLD general queries:3.
```

```
Received MLDv1 specific queries:0.
```

```
Received MLDv1 reports:12.
```

```
Received MLD dones:0.
```

```
Sent MLDv1 specific queries:0.
```

```
Received MLDv2 reports:0.
```

```
Received MLDv2 reports with right and wrong records:0.
```

```
Received MLDv2 specific queries:0.
```

```
Received MLDv2 specific sg queries:0.
```

```
Sent MLDv2 specific queries:0.
```

```
Sent MLDv2 specific sg queries:0.
```

```
Received error MLD messages:0.
```

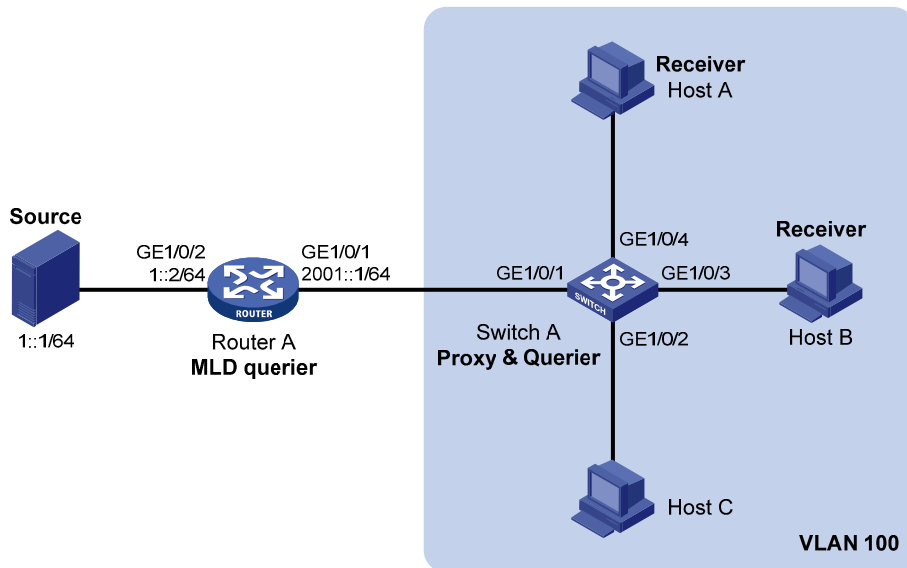
MLD snooping proxying configuration example

Network requirements

As shown in [Figure 71](#), MLDv1 runs on Router A and MLDv1 snooping runs on Switch A. Router A acts as the MLD querier.

Configure MLD snooping proxying on Switch A. This enables the switch to forward MLD reports and done messages on behalf of the attached hosts and to respond to MLD queries from Router A and then forward the queries to the hosts on behalf of Router A.

Figure 71 Network diagram



Configuration procedure

1. Configure IPv6 addresses for interfaces:
Configure an IP address and prefix length for each interface as per [Figure 71](#). (Details not shown.)
2. Configure Router A:
Enable IPv6 multicast routing, enable IPv6 PIM-DM on each interface, and enable MLD on port GigabitEthernet 1/0/1.

```
<RouterA> system-view  
[RouterA] multicast ipv6 routing-enable  
[RouterA] interface gigabitethernet 1/0/1  
[RouterA-GigabitEthernet1/0/1] mld enable  
[RouterA-GigabitEthernet1/0/1] pim ipv6 dm  
[RouterA-GigabitEthernet1/0/1] quit  
[RouterA] interface gigabitethernet 1/0/2  
[RouterA-GigabitEthernet1/0/2] pim ipv6 dm  
[RouterA-GigabitEthernet1/0/2] quit
```
3. Configure Switch A:
Enable MLD snooping globally.

```
<SwitchA> system-view  
[SwitchA] mld-snooping  
[SwitchA-mld-snooping] quit
```


Create VLAN 100, assign ports GigabitEthernet 1/0/1 through GigabitEthernet 1/0/4 to this VLAN, and enable MLD snooping and MLD snooping proxying in the VLAN.

```
[SwitchA] vlan 100  
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/4  
[SwitchA-vlan100] mld-snooping enable  
[SwitchA-vlan100] mld-snooping proxying enable  
[SwitchA-vlan100] quit
```
4. Verify the configuration:

After the configuration is completed, Host A and Host B send MLD join messages addressed to group FF1E::101. When receiving the messages, Switch A sends a join message for the group out of port GigabitEthernet 1/0/1 (a router port) to Router A. Use the **display mld-snooping group** command and the **display mld group** command to display information about MLD snooping groups and MLD multicast groups. For example:

Display information about MLD snooping groups on Switch A.

```
[SwitchA] display mld-snooping group
Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).

Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port
Subvlan flags: R-Real VLAN, C-Copy VLAN
Vlan(id):100.
Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).
Router port(s):total 1 port(s).
    GE1/0/1                (D)
IP group(s):the following ip group(s) match to one mac group.
IP group address:FF1E::101
(::, FF1E::101):
    Host port(s):total 2 port(s).
        GE1/0/3            (D)
        GE1/0/4            (D)
MAC group(s):
    MAC group address:3333-0000-0101
    Host port(s):total 2 port(s).
        GE1/0/3
        GE1/0/4
```

Display information about MLD multicast groups on Router A.

```
[RouterA] display mld group
Total 1 MLD Group(s).
Interface group report information
GigabitEthernet1/0/1(2001::1):
Total 1 MLD Group reported
Group Address: FF1E::1
Last Reporter: FE80::2FF:FFFF:FE00:1
Uptime: 00:00:03
Expires: 00:04:17
```

When Host A leaves the IPv6 multicast group, it sends an MLD done message to Switch A. Receiving the message, Switch A removes port GigabitEthernet 1/0/4 from the member port list of the forwarding entry for the group; however, it does not remove the group or forward the done message to Router A because Host B is still in the group. Use the **display mld-snooping group** command to display information about MLD snooping groups. For example:

Display information about MLD snooping groups on Switch A.

```
[SwitchA] display mld-snooping group
Total 1 IP Group(s).
```

```

Total 1 IP Source(s).
Total 1 MAC Group(s).
Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port
Subvlan flags: R-Real VLAN, C-Copy VLAN
Vlan(id):100.
    Total 1 IP Group(s).
    Total 1 IP Source(s).
    Total 1 MAC Group(s).
Router port(s):total 1 port(s).
    GE1/0/1                (D)
IP group(s):the following ip group(s) match to one mac group.
    IP group address:FF1E::101
    (:, FF1E::101):
        Host port(s):total 1 port(s).
        GE1/0/3            (D)
MAC group(s):
    MAC group address:3333-0000-0101
    Host port(s):total 1 port(s).
    GE1/0/3

```

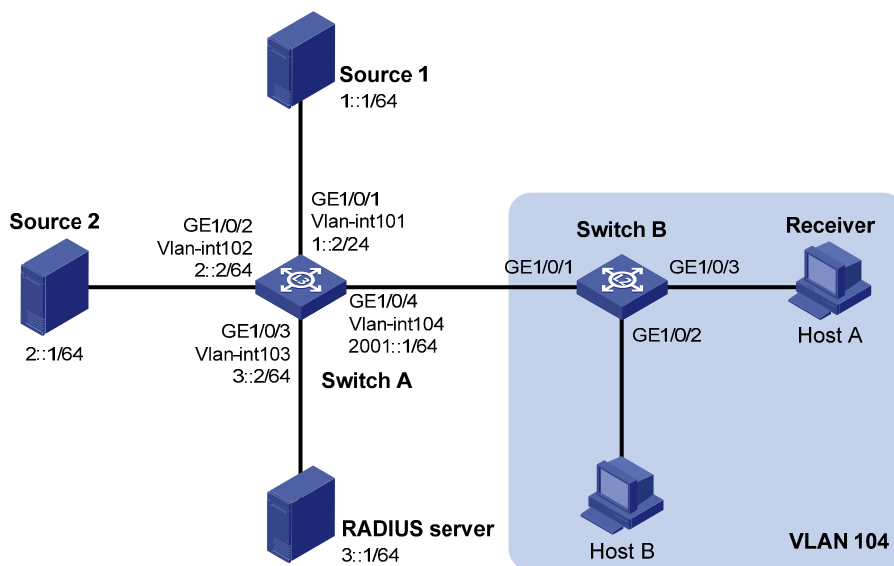
IPv6 multicast source and user control policy configuration example

Network requirements

As shown in [Figure 72](#), Switch A is a Layer-3 switch. MLDv1 runs on Switch A and MLDv1 snooping runs on Switch B. Multicast sources and hosts run 802.1X client.

An IPv6 multicast source control policy is configured on Switch A to block multicast flows from Source 2 to FF1E::101. An IPv6 multicast user control policy is configured on Switch B so that Host A can join or leave only multicast group FF1E::101.

Figure 72 Network diagram



Configuration procedures

1. Configure IP addresses for interfaces:

Enable IPv6 forwarding and configure an IP address and prefix length for each interface as per [Figure 72](#). (Details not shown.)

2. Configure Switch A:

Create VLAN 101 through VLAN 104 and assign GigabitEthernet 1/0/1 through GigabitEthernet 1/0/4 to the four VLANs respectively.

```
<SwitchA> system-view
[SwitchA] vlan 101
[SwitchA-vlan101] port gigabitethernet 1/0/1
[SwitchA-vlan101] quit
[SwitchA] vlan 102
[SwitchA-vlan102] port gigabitethernet 1/0/2
[SwitchA-vlan102] quit
[SwitchA] vlan 103
[SwitchA-vlan103] port gigabitethernet 1/0/3
[SwitchA-vlan103] quit
[SwitchA] vlan 104
[SwitchA-vlan104] port gigabitethernet 1/0/4
[SwitchA-vlan104] quit
```

Enable IPv6 multicast routing. Enable IPv6 PIM-DM on VLAN-interface 101, VLAN-interface 102 and VLAN-interface 104, and enable MLD on VLAN-interface 104.

```
[SwitchA] multicast ipv6 routing-enable
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim ipv6 dm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 102
[SwitchA-Vlan-interface102] pim ipv6 dm
[SwitchA-Vlan-interface102] quit
[SwitchA] interface vlan-interface 104
[SwitchA-Vlan-interface104] pim ipv6 dm
[SwitchA-Vlan-interface104] mld enable
[SwitchA-Vlan-interface104] quit
```

Create a multicast source control policy, **policy1**, so that multicast flows from Source 2 to FF1E::101 will be blocked.

```
[SwitchA] acl ipv6 number 3001
[SwitchA-acl6-adv-3001] rule permit udp source 2::1 128 destination ff1e::101 128
[SwitchA-acl6-adv-3001] quit
[SwitchA] traffic classifier classifier1
[SwitchA-classifier-classifier1] if-match acl ipv6 3001
[SwitchA-classifier-classifier1] quit
[SwitchA] traffic behavior behavior1
[SwitchA-behavior-behavior1] filter deny
[SwitchA-behavior-behavior1] quit
[SwitchA] qos policy policy1
[SwitchA-qospolicy-policy1] classifier classifier1 behavior behavior1
[SwitchA-qospolicy-policy1] quit
```

Create a user profile, apply **policy1** to the inbound direction of GE 1/0/2 in user profile view, and enable the user profile.

```
[SwitchA] user-profile profile1
[SwitchA-user-profile-profile1] qos apply policy policy1 inbound
[SwitchA-user-profile-profile1] quit
[SwitchA] user-profile profile1 enable
```

Create RADIUS scheme **scheme1**; set the service type for the RADIUS server to **extended**; specify the IP addresses of the primary authentication/authorization server and accounting server as 3::1; set the shared keys to 123321; specify that no domain name is carried in a username sent to the RADIUS server.

```
[SwitchA] radius scheme scheme1
[SwitchA-radius-scheme1] server-type extended
[SwitchA-radius-scheme1] primary authentication 3::1
[SwitchA-radius-scheme1] key authentication 123321
[SwitchA-radius-scheme1] primary accounting 3::1
[SwitchA-radius-scheme1] key accounting 123321
[SwitchA-radius-scheme1] user-name-format without-domain
[SwitchA-radius-scheme1] quit
```

Create an ISP domain **domain1**; reference **scheme1** for the authentication, authorization, and accounting for LAN users; specify **domain1** as the default ISP domain.

```
[SwitchA] domain domain1
[SwitchA-isp-domain1] authentication lan-access radius-scheme scheme1
[SwitchA-isp-domain1] authorization lan-access radius-scheme scheme1
[SwitchA-isp-domain1] accounting lan-access radius-scheme scheme1
[SwitchA-isp-domain1] quit
[SwitchA] domain default enable domain1
```

Globally enable 802.1X and then enable it on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2.

```
[SwitchA] dot1x
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] dot1x
[SwitchA-GigabitEthernet1/0/1] quit
[SwitchA] interface gigabitethernet 1/0/2
[SwitchA-GigabitEthernet1/0/2] dot1x
[SwitchA-GigabitEthernet1/0/2] quit
```

3. Configure Switch B:

Globally enable MLD snooping.

```
<SwitchB> system-view
[SwitchB] mld-snooping
[SwitchB-mld-snooping] quit
```

Create VLAN 104, assign GigabitEthernet 1/0/1 through GigabitEthernet 1/0/3 to this VLAN, and enable MLD snooping in this VLAN.

```
[SwitchB] vlan 104
[SwitchB-vlan104] port gigabitethernet 1/0/1 to gigabitethernet 1/0/3
[SwitchB-vlan104] mld-snooping enable
[SwitchB-vlan104] quit
```

Create a user profile **profile2** and configure the user profile so that users can join or leave only one IPv6 multicast group, FF1E::101. Then, enable the user profile.

```
[SwitchB] acl ipv6 number 2001
[SwitchB-acl6-basic-2001] rule permit source ff1e::101 128
[SwitchB-acl6-basic-2001] quit
[SwitchB] user-profile profile2
[SwitchB-user-profile-profile2] mld-snooping access-policy 2001
[SwitchB-user-profile-profile2] quit
[SwitchB] user-profile profile2 enable
```

Create a RADIUS scheme **scheme2**; set the service type for the RADIUS server to **extended**; specify the IP addresses of the primary authentication/authorization server and accounting server as 3::1; set the shared keys to 321123; specify that a username sent to the RADIUS server carry no domain name.

```
[SwitchB] radius scheme scheme2
[SwitchB-radius-scheme2] server-type extended
[SwitchB-radius-scheme2] primary authentication 3::1
[SwitchB-radius-scheme2] key authentication 321123
[SwitchB-radius-scheme2] primary accounting 3::1
[SwitchB-radius-scheme2] key accounting 321123
[SwitchB-radius-scheme2] user-name-format without-domain
[SwitchB-radius-scheme2] quit
```

Create an ISP domain **domain2**; reference **scheme2** for the authentication, authorization, and accounting for LAN users; specify **domain2** as the default ISP domain.

```
[SwitchB] domain domain2
[SwitchB-isp-domian2] authentication lan-access radius-scheme scheme2
[SwitchB-isp-domian2] authorization lan-access radius-scheme scheme2
[SwitchB-isp-domian2] accounting lan-access radius-scheme scheme2
[SwitchB-isp-domian2] quit
[SwitchB] domain default enable domain2
```

Globally enable 802.1X and then enable it on GigabitEthernet 1/0/2 and GigabitEthernet 1/0/3.

```
[SwitchB] dot1x
[SwitchB] interface gigabitethernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] dot1x
[SwitchB-GigabitEthernet1/0/2] quit
[SwitchB] interface gigabitethernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] dot1x
[SwitchB-GigabitEthernet1/0/3] quit
```

4. Configure RADIUS server:

On the RADIUS server, configure the parameters related to Switch A and Switch B. For more information, see the configuration guide of the RADIUS server.

5. Verify the configuration:

After the configurations, the two multicast sources and hosts initiate 802.1X authentication. After passing the authentication, Source 1 sends multicast flows to FF1E::101 and Source 2 sends multicast flows to FF1E::102; Host A sends report messages to join IPv6 multicast groups FF1E::101 and FF1E::102. Use the **display mld-snooping group** command to display information about MLD snooping groups. For example:

Display information about MLD snooping groups in VLAN 104 on Switch B.

```
[SwitchB] display mld-snooping group vlan 104 verbose
Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).

Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port
Subvlan flags: R-Real VLAN, C-Copy VLAN
Vlan(id):104.
Total 1 IP Group(s).
Total 1 IP Source(s).
Total 1 MAC Group(s).
Router port(s):total 1 port(s).
    GE1/0/1                (D) ( 00:01:30 )
IP group(s):the following ip group(s) match to one mac group.
  IP group address:FF1E::101
    (::, FF1E::101):
      Attribute:    Host Port
      Host port(s):total 1 port(s).
        GE1/0/3                (D) ( 00:04:10 )
MAC group(s):
  MAC group address:3333-0000-0101
  Host port(s):total 1 port(s).
    GE1/0/3
```

The output shows that GigabitEthernet 1/0/3 on Switch B has joined FF1E::101 but not FF1E::102.

Assume that Source 2 starts sending multicast traffic to FF1E::101. Use the **display multicast ipv6 forwarding-table** to display the IPv6 multicast forwarding table information.

Display the information about FF1E::101 in the IPv6 multicast forwarding table on Switch A.

```
[SwitchA] display multicast ipv6 forwarding-table ff1e::101
IPv6 Multicast Forwarding Table

Total 1 entry

Total 1 entry matched
00001. (1::1, FF1E::101)
  MID: 0, Flags: 0x0:0
  Uptime: 00:08:32, Timeout in: 00:03:26
  Incoming interface: Vlan-interface101
  List of 1 outgoing interfaces:
    1: Vlan-interface104
  Matched 19648 packets(20512512 bytes), Wrong If 0 packets
  Forwarded 19648 packets(20512512 bytes)
```

The output shows that Switch A maintains a multicast forwarding entry for multicast packets from Source 1 to FF1E::101. No forwarding entry exists for packets from Source 2 to FF1E::101, which indicates that IPv6 multicast packets from Source 2 are blocked.

Troubleshooting MLD snooping

Layer 2 multicast forwarding cannot function

Symptom

Layer 2 multicast forwarding cannot function.

Analysis

MLD snooping is not enabled.

Solution

1. Use the **display current-configuration** command to display the running status of MLD snooping.
2. If MLD snooping is not enabled, use the **mld-snooping** command to enable MLD snooping globally, and then use **mld-snooping enable** command to enable MLD snooping in VLAN view.
3. If MLD snooping is disabled only for the corresponding VLAN, use the **mld-snooping enable** command in VLAN view to enable MLD snooping in the corresponding VLAN.

Configured IPv6 multicast group policy fails to take effect

Symptom

Although an IPv6 multicast group policy has been configured to allow hosts to join specific IPv6 multicast groups, the hosts can still receive IPv6 multicast data addressed to other groups.

Analysis

- The IPv6 ACL rule is incorrectly configured.
- The IPv6 multicast group policy is not correctly applied.
- The function of dropping unknown IPv6 multicast data is not enabled, so unknown IPv6 multicast data is flooded.

Solution

1. Use the **display acl ipv6** command to check the configured IPv6 ACL rule. Make sure that the IPv6 ACL rule conforms to the IPv6 multicast group policy to be implemented.
2. Use the **display this** command in MLD-snooping view or the corresponding interface view to verify that the correct IPv6 multicast group policy has been applied. If not, use the **group-policy** or **mld-snooping group-policy** command to apply the correct IPv6 multicast group policy.
3. Use the **display current-configuration** command to verify that the function of dropping unknown IPv6 multicast data is enabled. If not, use the **drop-unknown** or **mld-snooping drop-unknown** command to enable the function of dropping unknown IPv6 multicast data.

Appendix

Processing of IPv6 multicast protocol messages

With Layer 3 multicast routing enabled, an MLD snooping-enabled switch processes IPv6 multicast protocol messages differently under different conditions, as follows:

1. If only MLD is enabled on the switch, or if both MLD and IPv6 PIM are enabled on the switch, the switch does the following:
 - Maintains dynamic member ports or dynamic router ports according to MLD packets
 - Maintains dynamic router ports according to IPv6 PIM hello packets
2. If only IPv6 PIM is enabled on the switch, the following occurs:
 - The switch broadcasts MLD messages as unknown messages in the VLAN.
 - After receiving an IPv6 PIM hello message, the switch maintains the corresponding dynamic router port.
3. If MLD is disabled on the switch, one of the following occurs:
 - If IPv6 PIM is disabled, the switch deletes all its dynamic member ports and dynamic router ports.
 - If IPv6 PIM is enabled, the switch deletes only its dynamic member ports but not its dynamic router ports.

NOTE:

On a switch with Layer-3 IPv6 multicast routing enabled, use the **display mld group port-info** command to display Layer-2 port information. For more information about the **display mld group port-info** command, see *IP Multicast Command Reference*.

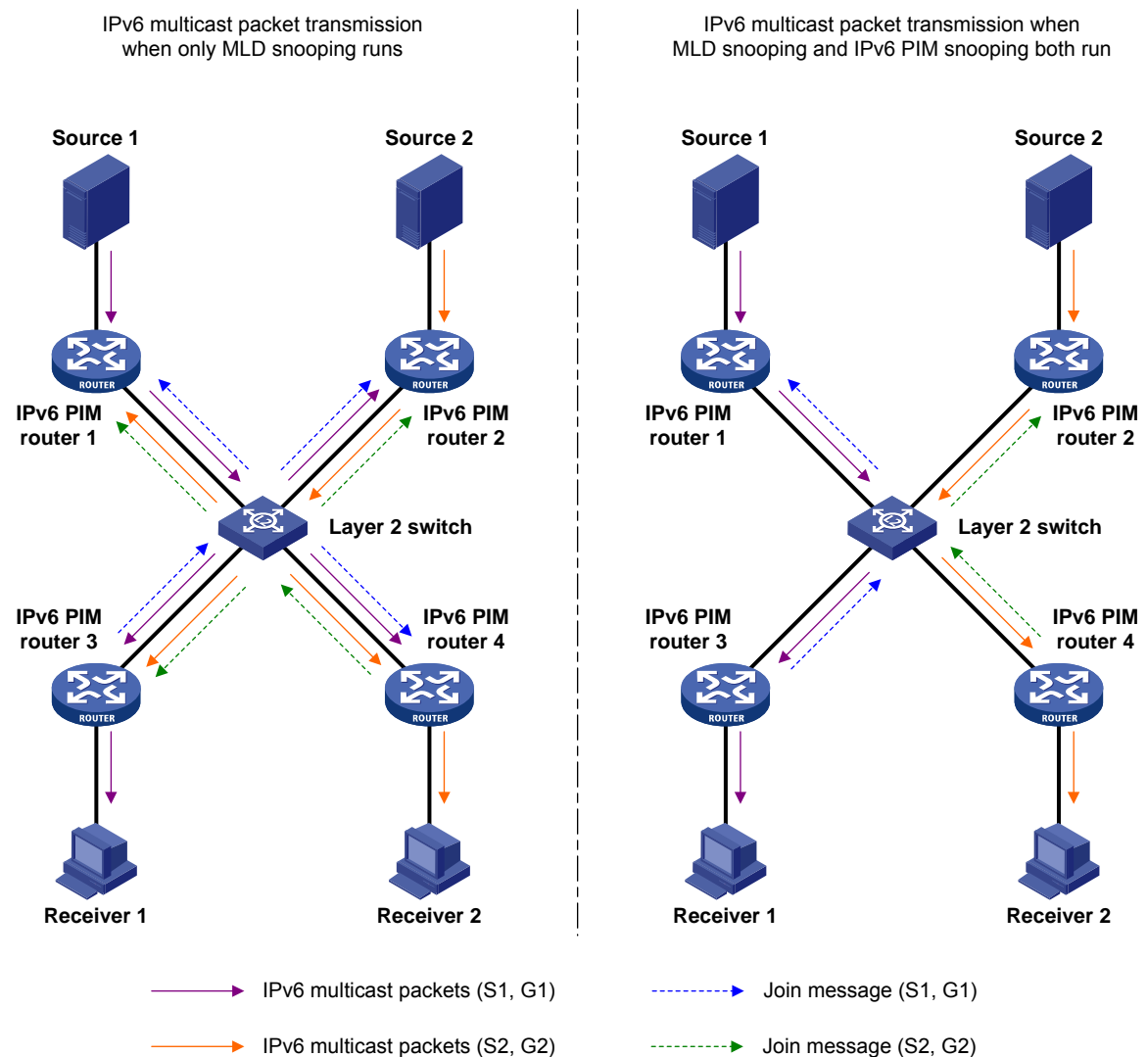
4. If IPv6 PIM is disabled on the switch, one of the following occurs:
 - If MLD is disabled, the switch deletes all its dynamic router ports.
 - If MLD is enabled, the switch maintains all its dynamic member ports and dynamic router ports.

Configuring IPv6 PIM snooping

Overview

IPv6 Protocol Independent Multicast (PIM) snooping runs on Layer 2 devices. It determines which ports are interested in multicast data by analyzing the received IPv6 PIM messages, and adds the ports to a multicast forwarding entry to make sure that multicast data can be forwarded to only the ports that are interested in the data.

Figure 73 Multicast packet transmission without or with IPv6 PIM snooping



As shown in Figure 73, Source 1 sends multicast data to multicast group G1, and Source 2 sends multicast data to multicast group G2. Receiver 1 belongs to G1, and Receiver 2 belongs to G2. The Layer 2 switch's interfaces that connect to the IPv6 PIM-capable routers are in the same VLAN.

- When running MLD snooping without IPv6 PIM snooping, the Layer 2 switch maintains the router ports according to IPv6 PIM hello messages received from IPv6 PIM-capable routers, broadcasts all

other types of received IPv6 PIM messages in the VLAN, and forwards all multicast data to all router ports in the VLAN. Each IPv6 PIM-capable router in the VLAN, whether interested in the multicast data or not, will receive all multicast data and all IPv6 PIM messages except for IPv6 PIM hello messages.

- If the Layer 2 switch runs both MLD snooping and IPv6 PIM snooping, it determines whether an IPv6 PIM-capable router is interested in the multicast data destined for a multicast group according to the received IPv6 PIM messages that the router sends, and adds the port that connects to the router to a multicast forwarding entry. Then, the Layer 2 switch can correctly forward IPv6 PIM messages and the multicast data only to the router according to the multicast forwarding entry, saving network bandwidth.

For more information about MLD snooping and the router port, see "[Configuring MLD snooping](#)."

For more information about IPv6 PIM, see "[Configuring IPv6 PIM \(available only on the HP 5500 EI\)](#)"

Configuring IPv6 PIM snooping

Configuration guidelines

Before you configure IPv6 PIM snooping for a VLAN, you must enable IPv6 forwarding and MLD snooping globally and enable MLD snooping in the VLAN.

IPv6 PIM snooping does not work in the sub-VLANs of a multicast VLAN. For more information about IPv6 multicast VLAN, see "[Configuring IPv6 multicast VLANs](#)."

In a network with IPv6 PIM snooping enabled switches, configure the size of each join/prune message no more than the path maximum transmission unit (MTU) on the IPv6 PIM-enabled edge router on the receiver side. For more information about the join/prune messages, see "[Configuring IPv6 PIM \(available only on the HP 5500 EI\)](#)."

After you enable IPv6 PIM snooping in a VLAN, IPv6 PIM snooping works only on the member interfaces of the VLAN.

Configuration procedure

To configure IPv6 PIM snooping:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IPv6 forwarding globally.	ipv6	Disabled by default
3. Enable MLD snooping globally and enter MLD-snooping view.	mld-snooping	Disabled by default
4. Return to system view.	quit	N/A
5. Enter VLAN view.	vlan <i>vlan-id</i>	N/A
6. Enable MLD snooping in the VLAN	mld-snooping enable	Disabled by default
7. Enable IPv6 PIM snooping in the VLAN	pim-snooping ipv6 enable	Disabled by default

For more information about the **mld-snooping** and **mld-snooping enable** commands, see *IP Multicast Command Reference*.

Displaying and maintaining IPv6 PIM snooping

Task	Command	Remarks
Display IPv6 PIM snooping neighbor information.	display pim-snooping ipv6 neighbor [vlan <i>vlan-id</i>] [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 PIM snooping routing entries.	display pim-snooping ipv6 routing-table [vlan <i>vlan-id</i>] [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the statistics information of IPv6 PIM messages learned by IPv6 PIM snooping..	display pim-snooping ipv6 statistics [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear the statistics information of IPv6 PIM messages learned by IPv6 PIM snooping..	reset pim-snooping ipv6 statistics	Available in user view

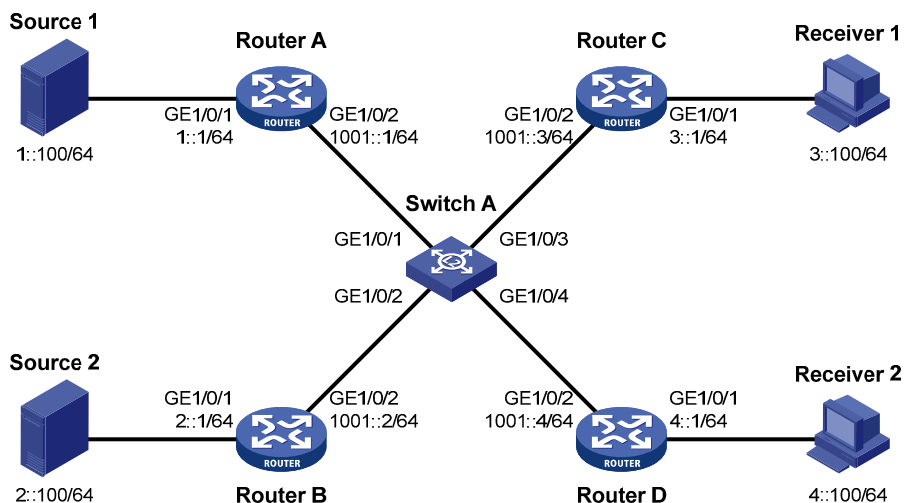
IPv6 PIM snooping configuration example

Network requirements

As shown in Figure 74, Source 1 sends multicast data to IPv6 multicast group FF1E::101, and Source 2 sends multicast data to IPv6 multicast group FF2E::101. Receiver 1 belongs to multicast group FF1E::101, and Receiver 2 belongs to multicast group FF2E::101. Router C and Router D run MLD on their interface GigabitEthernet 1/0/1. Router A, Router B, Router C, and Router D run IPv6 PIM-SM, and interface GigabitEthernet 1/0/2 on Router A acts as a C-BSR and C-RP.

Configure MLD snooping and IPv6 PIM snooping on Switch A so that Switch A forwards IPv6 PIM messages and multicast data to only the routers that are interested in the multicast data.

Figure 74 Network diagram



Configuration procedure

1. Enable IPv6 forwarding, and assign IPv6 addresses:

Enable IPv6 forwarding on the devices, configure an IPv6 address and prefix length for each interface according to [Figure 74](#). (Details not shown.)

2. Configure Router A:

Enable IPv6 multicast routing, enable IPv6 PIM-SM on each interface, and configure interface GigabitEthernet 1/0/2 as a C-BSR and C-RP.

```
<RouterA> system-view
[RouterA] multicast ipv6 routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] pim ipv6 sm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] pim ipv6 sm
[RouterA-GigabitEthernet1/0/2] quit
[RouterA] pim ipv6
[RouterA-pim6] c-bsr 1001::1
[RouterA-pim6] c-rp 1001::1
```

3. Configure Router B:

Enable IPv6 multicast routing, and enable IPv6 PIM-SM on each interface.

```
<RouterB> system-view
[RouterB] multicast ipv6 routing-enable
[RouterB] interface gigabitethernet 1/0/1
[RouterB-GigabitEthernet1/0/1] pim ipv6 sm
[RouterB-GigabitEthernet1/0/1] quit
[RouterB] interface gigabitethernet 1/0/2
[RouterB-GigabitEthernet1/0/2] pim ipv6 sm
```

4. Configure Router C:

Enable IPv6 multicast routing, enable IPv6 PIM-SM on each interface, and enable MLD on GigabitEthernet 1/0/1.

```
<RouterC> system-view
[RouterC] multicast ipv6 routing-enable
[RouterC] interface gigabitethernet 1/0/1
[RouterC-GigabitEthernet1/0/1] pim ipv6 sm
[RouterC-GigabitEthernet1/0/1] mld enable
[RouterC-GigabitEthernet1/0/1] quit
[RouterC] interface gigabitethernet 1/0/2
[RouterC-GigabitEthernet1/0/2] pim ipv6 sm
```

5. Configure Router D:

The configuration on Router D is similar to that on Router C. (Details not shown.)

6. Configure Switch A:

Enable MLD snooping globally.

```
<SwitchA> system-view
[SwitchA] mld-snooping
[SwitchA-mld-snooping] quit
```

Create VLAN 100, assign GigabitEthernet 1/0/1 through GigabitEthernet 1/0/4 to this VLAN, and enable MLD snooping and IPv6 PIM snooping in the VLAN.

```
[SwitchA] vlan 100
[SwitchA-vlan100] port gigabitethernet 1/0/1 to gigabitethernet 1/0/4
[SwitchA-vlan100] mld-snooping enable
[SwitchA-vlan100] pim-snooping ipv6 enable
[SwitchA-vlan100] quit
```

7. Verify the configuration:

On Switch A, display the IPv6 PIM snooping neighbor information of VLAN 100.

```
[SwitchA] display pim-snooping ipv6 neighbor vlan 100
Total number of neighbors: 4
```

```
VLAN ID: 100
Total number of neighbors: 4
Neighbor      Port          Expires      Option Flags
FE80::1       GE1/0/1       02:02:23    LAN Prune Delay
FE80::2       GE1/0/2       03:00:05    LAN Prune Delay
FE80::3       GE1/0/3       02:22:13    LAN Prune Delay
FE80::4       GE1/0/4       03:07:22    LAN Prune Delay
```

The output shows that Router A, Router B, Router C, and Router D are IPv6 PIM snooping neighbors.

On Switch A, display the IPv6 PIM snooping routing information of VLAN 100.

```
[SwitchA] display pim-snooping ipv6 routing-table vlan 100 slot 1
Total 2 entry(ies)
FSM Flag: NI-no info, J-join, PP-prune pending
```

```
VLAN ID: 100
Total 2 entry(ies)
(*, FF1E::101)
  Upstream neighbor: FE80::1
  Upstream port: GE1/0/1
  Total number of downstream ports: 1
  1: GE1/0/3
  Expires: 00:03:01, FSM: J
(*, FF2E::101)
  Upstream neighbor: FE80::1
  Upstream port: GE1/0/2
  Total number of downstream ports: 1
  1: GE1/0/4
  Expires: 00:01:05, FSM: J
```

The output shows that Switch A will forward the multicast data intended for IPv6 multicast group FF1E::101 to only Router C, and forward the multicast data intended for IPv6 multicast group FF2E::101 to only Router D.

Troubleshooting IPv6 PIM snooping

IPv6 PIM snooping does not work

Symptom

IPv6 PIM snooping does not work.

Analysis

MLD snooping or IPv6 PIM snooping is not enabled on the switch.

Solution

1. Use the **display current-configuration** command to check the status of MLD snooping and IPv6 PIM snooping.
2. If MLD snooping is not enabled, enter system view and use the **mld-snooping** command to enable MLD snooping globally. Then, enter VLAN view and use the **mld-snooping enable** and **pim-snooping ipv6 enable** commands to enable MLD snooping and IPv6 PIM snooping in the VLAN.
3. If IPv6 PIM snooping is not enabled, enter VLAN view and use the **pim-snooping ipv6 enable** command to enable IPv6 PIM snooping in the VLAN.

Some downstream IPv6 PIM-capable routers cannot receive multicast data

Symptom

In a network with fragmented join/prune messages, some downstream IPv6 PIM-capable routers cannot receive multicast data.

Analysis

IPv6 PIM snooping cannot reassemble messages, and it cannot maintain the status of downstream routers that the join/prune message fragments carry. To ensure the normal operation of the system, IPv6 PIM snooping must broadcast join/prune message fragments in the VLAN. However, if the VLAN has an IPv6 PIM-capable router that has the join suppression function enabled, the broadcast join/prune message fragments might suppress the join messages of other IPv6 PIM-capable routers in the VLAN. As a result, some IPv6 PIM-capable routers cannot receive the multicast data addressed to a specific multicast group because their join messages are suppressed. To solve this problem, disable the join suppression function on all IPv6 PIM-capable routers that connect to the IPv6 PIM snooping-capable switch in the VLAN.

Solution

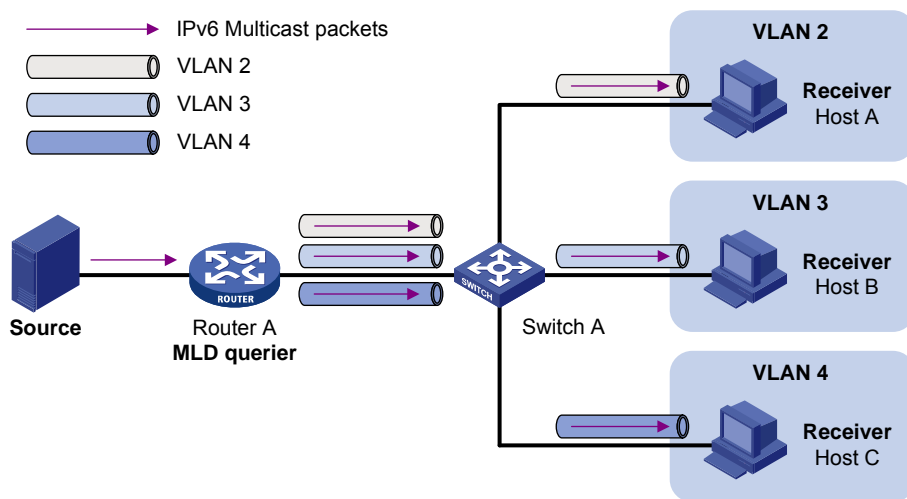
1. Use the **pim ipv6 hello-option neighbor-tracking** command to enable the neighbor tracking function on the interfaces of IPv6 PIM-capable routers that connect to the IPv6 PIM snooping-capable switch.
2. If the network has an IPv6 PIM-capable router that cannot be enabled with the neighbor tracking function, be sure to disable IPv6 PIM snooping on the IPv6 PIM snooping-capable switch.

Configuring IPv6 multicast VLANs

Overview

As shown in [Figure 75](#), in the traditional IPv6 multicast programs-on-demand mode, when hosts (Host A, Host B, and Host C), which belong to different VLANs, require IPv6 multicast programs on demand service, the Layer 3 device, Router A, must forward a separate copy of the multicast traffic in each user VLAN to the Layer 2 device, Switch A. This results in not only waste of network bandwidth but also extra burden on the Layer 3 device.

Figure 75 Multicast transmission without IPv6 multicast VLAN



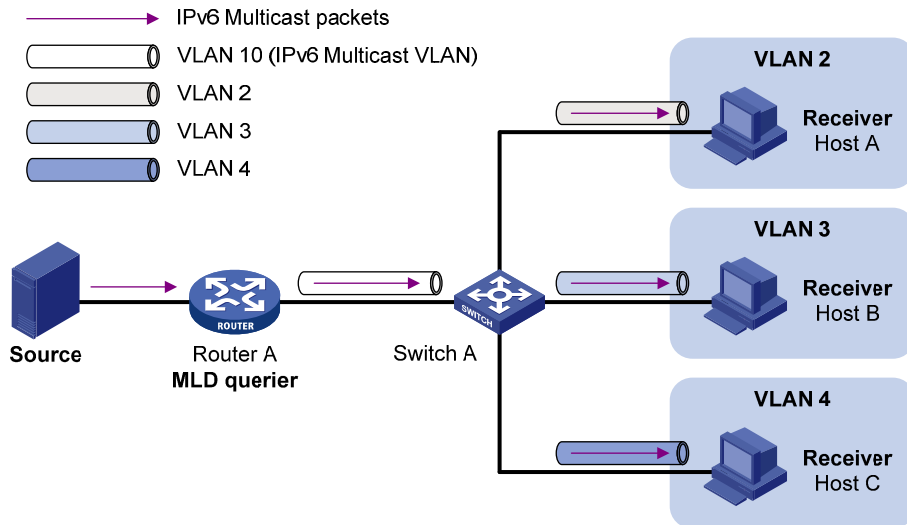
The IPv6 multicast VLAN feature configured on the Layer 2 device is the solution to this issue. With the IPv6 multicast VLAN feature, the Layer 3 device needs to replicate the multicast traffic only in the IPv6 multicast VLAN instead of making a separate copy of the multicast traffic in each user VLAN. This saves the network bandwidth and lessens the burden of the Layer 3 device.

The IPv6 multicast VLAN feature can be implemented in sub-VLAN-based IPv6 multicast VLAN and port-based IPv6 multicast VLAN.

Sub-VLAN-based IPv6 multicast VLAN

As shown in [Figure 76](#), Host A, Host B and Host C are in different user VLANs. On Switch A, configure VLAN 10 as an IPv6 multicast VLAN, configure all the user VLANs as sub-VLANs of VLAN 10, and enable MLD snooping in the IPv6 multicast VLAN.

Figure 76 Sub-VLAN-based IPv6 multicast VLAN

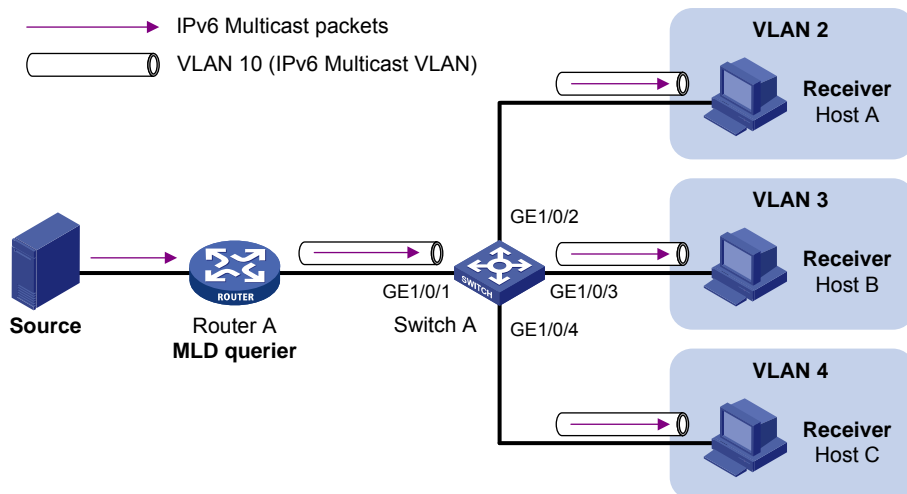


After the configuration, MLD snooping manages router ports in the IPv6 multicast VLAN and member ports in the sub-VLANs. When forwarding multicast data to Switch A, Router A sends only one copy of multicast data to Switch A in the IPv6 multicast VLAN, and Switch A distributes the data to the sub-VLANs that contain receivers.

Port-based IPv6 multicast VLAN

As shown in Figure 77, Host A, Host B, and Host C are in different user VLANs. All the user ports are hybrid ports. On Switch A, configure VLAN 10 as an IPv6 multicast VLAN, assign all the user ports to VLAN 10, and enable MLD snooping in the IPv6 multicast VLAN and all user VLANs.

Figure 77 Port-based IPv6 multicast VLAN



After the configuration, if Switch A receives an MLD message on a user port, it tags the message with the IPv6 multicast VLAN ID and relays it to the MLD querier, so that MLD snooping can uniformly manage the router ports and member ports in the IPv6 multicast VLAN. When Router A forwards multicast data to Switch A, it sends only one copy of multicast data to Switch A in the IPv6 multicast VLAN, and Switch A distributes the data to all member ports in the IPv6 multicast VLAN.

For more information about MLD snooping, router ports, and member ports, see "[Configuring MLD snooping.](#)"

For more information about VLAN tags, see *Layer 2—LAN Switching Configuration Guide*.

IPv6 multicast VLAN configuration task list

Configuration task	Remarks
Configuring a sub-VLAN-based IPv6 multicast VLAN	Required
Configuring a port-based IPv6 multicast VLAN	Configuring user port attributes Configuring IPv6 multicast VLAN ports

NOTE:

If you have configured both sub-VLAN-based IPv6 multicast VLAN and port-based IPv6 multicast VLAN on a device, the port-based IPv6 multicast VLAN configuration is given preference.

Configuring a sub-VLAN-based IPv6 multicast VLAN

Configuration prerequisites

Before you configure a sub-VLAN-based IPv6 multicast VLAN, complete the following tasks:

- Create VLANs as required.
- Enable MLD snooping in the VLAN to be configured as an IPv6 multicast VLAN.

Configuration guidelines

- You cannot configure an IPv6 multicast VLAN on a device with IP multicast routing enabled.
- The VLAN to be configured as an IPv6 multicast VLAN must exist.
- The VLANs to be configured as the sub-VLANs of the IPv6 multicast VLAN must exist and must not be IPv6 multicast VLANs or sub-VLANs of any other IPv6 multicast VLAN.
- The total number of sub-VLANs of an IPv6 multicast VLAN must not exceed the maximum number the system can support.

Configuration procedure

In this approach, you configure a VLAN as an IPv6 multicast VLAN, and configure user VLANs as sub-VLANs of the IPv6 multicast VLAN.

To configure a sub-VLAN-based IPv6 multicast VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the specified VLAN as an IPv6 multicast VLAN and enter IPv6 multicast VLAN view.	multicast-vlan ipv6 <i>vlan-id</i>	No IPv6 multicast VLAN configured by default.

Step	Command	Remarks
3. Configure the specified VLANs as sub-VLANs of the IPv6 multicast VLAN.	subvlan <i>vlan-list</i>	By default, an IPv6 multicast VLAN has no sub-VLANs.

Configuring a port-based IPv6 multicast VLAN

When you configure a port-based IPv6 multicast VLAN, you need to configure the attributes of each user port and then assign the ports to the IPv6 multicast VLAN.

A user port can be configured as a multicast VLAN port only if it is an Ethernet port or Layer 2 aggregate interface.

In Ethernet interface view or Layer 2 aggregate interface view, configurations that you make are effective only on the current interface. In port group view, configurations that you make are effective on all ports in the current port group.

Configuration prerequisites

Before you configure a port-based IPv6 multicast VLAN, complete the following tasks:

- Create VLANs as required.
- Enable MLD snooping in the VLAN to be configured as an IPv6 multicast VLAN.
- Enable MLD snooping in all the user VLANs.

Configuring user port attributes

Configure the user ports as hybrid ports to permit packets of the specified user VLAN to pass and configure the user VLAN to which the user ports belong as the default VLAN.

Configure the user ports to permit packets of the IPv6 multicast VLAN to pass and untag the packets. After receiving multicast packets tagged with the IPv6 multicast VLAN ID from the upstream device, the Layer 2 device untags the multicast packets and forwards them to its downstream device.

To configure user port attributes:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> • Enter interface view: interface <i>interface-type</i> <i>interface-number</i> • Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.
3. Configure the user port link type as hybrid.	port link-type hybrid	Access by default.
4. Specify the user VLAN that comprises the current user ports as the default VLAN.	port hybrid pvid vlan <i>vlan-id</i>	VLAN 1 by default.

Step	Command	Remarks
5. Configure the current user ports to permit packets of the specified IPv6 multicast VLAN to pass and untag the packets.	port hybrid vlan <i>vlan-id-list</i> { tagged untagged }	By default, a hybrid port permits only packets of VLAN 1 to pass.

For more information about the **port link-type**, **port hybrid pvid vlan**, and **port hybrid vlan** commands, see *Layer 2—LAN Switching Command Reference*.

Configuring IPv6 multicast VLAN ports

Configuration guidelines

In this approach, you configure a VLAN as an IPv6 multicast VLAN and assign user ports to it. You can do this by either adding the user ports in the IPv6 multicast VLAN or specifying the IPv6 multicast VLAN on the user ports. These two methods provide the same result.

You cannot configure an IPv6 multicast VLAN on a device with multicast routing enabled.

The VLAN to be configured as an IPv6 multicast VLAN must exist.

A port can belong to only one IPv6 multicast VLAN.

Configuration procedure

To configure IPv6 multicast VLAN ports in IPv6 multicast VLAN view:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the specified VLAN as an IPv6 multicast VLAN and enter IPv6 multicast VLAN view.	multicast-vlan ipv6 <i>vlan-id</i>	No IPv6 multicast VLAN configured by default.
3. Configure the ports as member ports of the IPv6 multicast VLAN.	port <i>interface-list</i>	By default, an IPv6 multicast VLAN has no member ports.

To configure IPv6 multicast VLAN ports in interface view or port group view:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the specified VLAN as an IPv6 multicast VLAN and enter IPv6 multicast VLAN view.	multicast-vlan ipv6 <i>vlan-id</i>	Not an IPv6 multicast VLAN by default.
3. Return to system view.	quit	N/A
4. Enter interface view or port group view.	<ul style="list-style-type: none"> • Enter interface view: interface <i>interface-type interface-number</i> • Enter port group view: port-group manual <i>port-group-name</i> 	Use either command.

Step	Command	Remarks
5. Configure the ports as member ports of the IPv6 multicast VLAN.	port multicast-vlan ipv6 <i>vlan-id</i>	By default, a user port does not belong to any IPv6 multicast VLAN.

Displaying and maintaining IPv6 multicast VLAN

Task	Command	Remarks
Display information about an IPv6 multicast VLAN.	display multicast-vlan ipv6 [<i>vlan-id</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

IPv6 multicast VLAN configuration examples

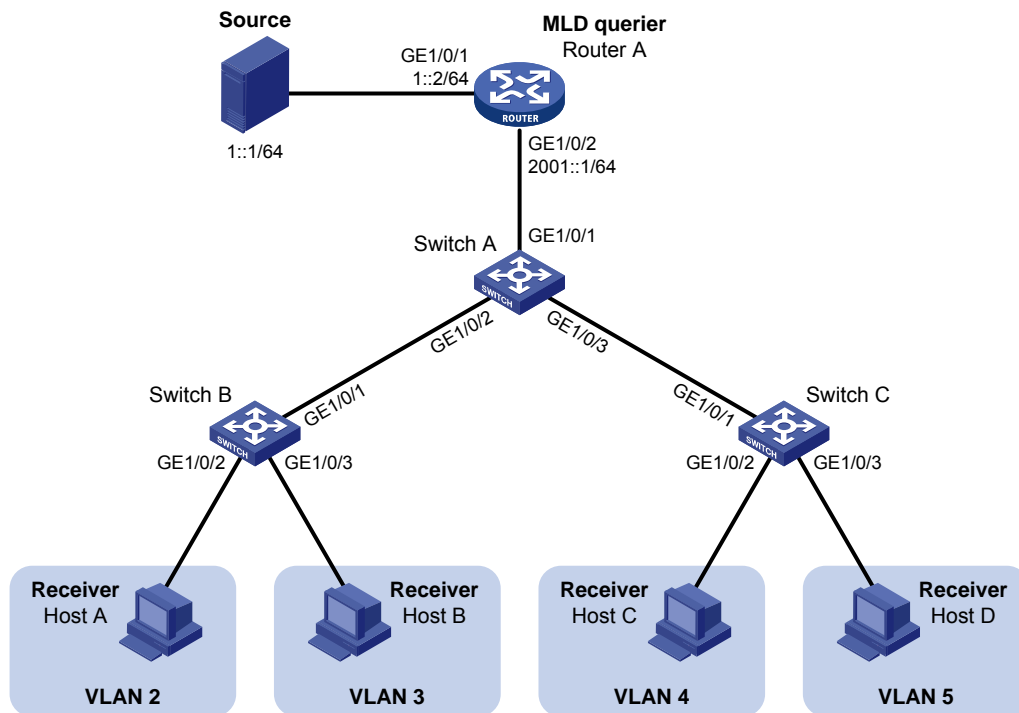
Sub-VLAN-based multicast VLAN configuration example

Network requirements

As shown in [Figure 78](#), MLDv1 runs on Router A, and MLDv1 snooping runs on Switch A. Router A acts as the MLD querier. The IPv6 multicast source sends IPv6 multicast data to the IPv6 multicast group FF1E::101. Host A, Host B, Host C, and Host D are receivers of the IPv6 multicast group. The hosts belong to VLAN 2 through VLAN 5 respectively.

Configure the sub-VLAN-based IPv6 multicast VLAN feature on Switch A so that Router A just sends IPv6 multicast data to Switch A through the IPv6 multicast VLAN and Switch A forwards the traffic to the receivers that belong to different user VLANs.

Figure 78 Network diagram



Configuration procedure

1. Enable IPv6 forwarding and configure IPv6 addresses:
Enable IPv6 forwarding on each device and configure an IPv6 address and address prefix for each interface as per Figure 78. (Details not shown.)
2. Configure Router A:
Enable IPv6 multicast routing, enable IPv6 PIM-DM on each interface and enable MLD on the host-side interface GigabitEthernet 1/0/2.

```
<RouterA> system-view  
[RouterA] multicast ipv6 routing-enable  
[RouterA] interface gigabitethernet 1/0/1  
[RouterA-GigabitEthernet1/0/1] pim ipv6 dm  
[RouterA-GigabitEthernet1/0/1] quit  
[RouterA] interface gigabitethernet 1/0/2  
[RouterA-GigabitEthernet1/0/2] pim ipv6 dm  
[RouterA-GigabitEthernet1/0/2] mld enable
```
3. Configure Switch A:
Enable MLD snooping globally.

```
<SwitchA> system-view  
[SwitchA] mld-snooping  
[SwitchA-mld-snooping] quit
```


Create VLAN 2 through VLAN 5.

```
[SwitchA] vlan 2 to 5
```


Configure GigabitEthernet 1/0/2 as a trunk port that permits packets from VLAN 2 and VLAN 3 to pass through.

```
[SwitchA] interface gigabitethernet 1/0/2
```

```
[SwitchA-GigabitEthernet1/0/2] port link-type trunk
[SwitchA-GigabitEthernet1/0/2] port trunk permit vlan 2 3
[SwitchA-GigabitEthernet1/0/2] quit
# Configure GigabitEthernet 1/0/3 as a trunk port that permits packets from VLAN 4 and VLAN 5 to pass through.
```

```
[SwitchA] interface gigabitethernet 1/0/3
[SwitchA-GigabitEthernet1/0/3] port link-type trunk
[SwitchA-GigabitEthernet1/0/3] port trunk permit vlan 4 5
[SwitchA-GigabitEthernet1/0/3] quit
```

Create VLAN 10, assign GigabitEthernet 1/0/1 to this VLAN and enable MLD snooping in the VLAN.

```
[SwitchA] vlan 10
[SwitchA-vlan10] port gigabitethernet 1/0/1
[SwitchA-vlan10] mld-snooping enable
[SwitchA-vlan10] quit
```

Configure VLAN 10 as an IPv6 multicast VLAN and configure VLAN 2 through VLAN 5 as its sub-VLANs.

```
[SwitchA] multicast-vlan ipv6 10
[SwitchA-ipv6-mvlan-10] subvlan 2 to 5
[SwitchA-ipv6-mvlan-10] quit
```

4. Configure Switch B:

Enable MLD snooping globally.

```
<SwitchB> system-view
[SwitchB] mld-snooping
[SwitchB-mld-snooping] quit
```

Create VLAN 2, assign GigabitEthernet 1/0/2 to VLAN 2, and enable MLD snooping in the VLAN.

```
[SwitchB] vlan 2
[SwitchB-vlan2] port gigabitethernet 1/0/2
[SwitchB-vlan2] mld-snooping enable
[SwitchB-vlan2] quit
```

Create VLAN 3, assign GigabitEthernet 1/0/3 to VLAN 3, and enable MLD snooping in the VLAN.

```
[SwitchB] vlan 3
[SwitchB-vlan3] port gigabitethernet 1/0/3
[SwitchB-vlan3] mld-snooping enable
[SwitchB-vlan3] quit
```

Configure GigabitEthernet 1/0/1 as a trunk port that permits packets from VLAN 2 and VLAN 3 to pass through.

```
[SwitchB] interface gigabitethernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] port link-type trunk
[SwitchB-GigabitEthernet1/0/1] port trunk permit vlan 2 3
```

5. Configure Switch C:

The configurations required on Switch C are similar to those on Switch B.

6. Verify the configuration:

Display information about the IPv6 multicast VLAN.

```
[SwitchA] display multicast-vlan ipv6
```

```
Total 1 IPv6 multicast-vlan(s)
```

```
IPv6 Multicast vlan 10
```

```
  subvlan list:
```

```
    vlan 2-5
```

```
  port list:
```

```
    no port
```

View the MLD snooping IPv6 multicast group information on Switch A.

```
[SwitchA] display mld-snooping group
```

```
Total 5 IP Group(s).
```

```
Total 5 IP Source(s).
```

```
Total 5 MAC Group(s).
```

```
Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port
```

```
Subvlan flags: R-Real VLAN, C-Copy VLAN
```

```
Vlan(id):2.
```

```
  Total 1 IP Group(s).
```

```
  Total 1 IP Source(s).
```

```
  Total 1 MAC Group(s).
```

```
  Router port(s):total 0 port(s).
```

```
  IP group(s):the following ip group(s) match to one mac group.
```

```
    IP group address:FF1E::101
```

```
      (::, FF1E::101):
```

```
        Host port(s):total 1 port(s).
```

```
          GE1/0/2          (D)
```

```
  MAC group(s):
```

```
    MAC group address:3333-0000-0101
```

```
      Host port(s):total 1 port(s).
```

```
        GE1/0/2
```

```
Vlan(id):3.
```

```
  Total 1 IP Group(s).
```

```
  Total 1 IP Source(s).
```

```
  Total 1 MAC Group(s).
```

```
  Router port(s):total 0 port(s).
```

```
  IP group(s):the following ip group(s) match to one mac group.
```

```
    IP group address:FF1E::101
```

```
      (::, FF1E::101):
```

```
        Host port(s):total 1 port(s).
```

```
          GE1/0/2          (D)
```

```
  MAC group(s):
```

```
    MAC group address:3333-0000-0101
```

```
      Host port(s):total 1 port(s).
```

```
        GE1/0/2
```

```
Vlan(id):4.
```

```
  Total 1 IP Group(s).
```

```
  Total 1 IP Source(s).
```

```
  Total 1 MAC Group(s).
```

```
  Router port(s):total 0 port(s).
```

```
  IP group(s):the following ip group(s) match to one mac group.
```

```

IP group address:FF1E::101
  (::, FF1E::101):
    Host port(s):total 1 port(s).
      GE1/0/3                (D)
MAC group(s):
  MAC group address:3333-0000-0101
    Host port(s):total 1 port(s).
      GE1/0/3
Vlan(id):5.
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
  Router port(s):total 0 port(s).
  IP group(s):the following ip group(s) match to one mac group.
    IP group address:FF1E::101
      (::, FF1E::101):
        Host port(s):total 1 port(s).
          GE1/0/3            (D)
    MAC group(s):
      MAC group address:3333-0000-0101
        Host port(s):total 1 port(s).
          GE1/0/3
Vlan(id):10.
  Total 1 IP Group(s).
  Total 1 IP Source(s).
  Total 1 MAC Group(s).
  Router port(s):total 1 port(s).
    GE1/0/1                (D)
  IP group(s):the following ip group(s) match to one mac group.
    IP group address:FF1E::101
      (::, FF1E::101):
        Host port(s):total 0 port(s).
    MAC group(s):
      MAC group address:3333-0000-0101
        Host port(s):total 0 port(s).

```

The output shows that MLD snooping is maintaining the router port in the IPv6 multicast VLAN (VLAN 10) and the member ports in the sub-VLANs (VLAN 2 through VLAN 5).

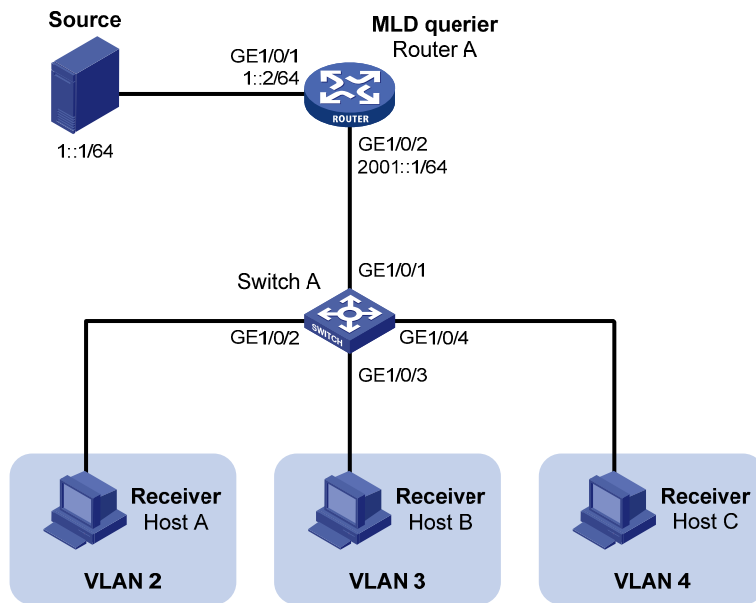
Port-based multicast VLAN configuration example

Network requirements

As shown in [Figure 79](#), MLDv1 runs on Router A. MLDv1 snooping runs on Switch A. Router A acts as the MLD querier. The IPv6 multicast source sends IPv6 multicast data to IPv6 multicast group FF1E::101. Host A, Host B, and Host C are receivers of the IPv6 multicast group. The hosts belong to VLAN 2 through VLAN 4 respectively.

Configure the port-based IPv6 multicast VLAN feature on Switch A so that Router A sends IPv6 multicast data to Switch A through the IPv6 multicast VLAN, and Switch A forwards the IPv6 multicast data to the receivers that belong to different user VLANs.

Figure 79 Network diagram



Configuration procedure

1. Enable IPv6 forwarding, and configure IPv6 addresses:
Enable IPv6 forwarding on each device, and configure the IPv6 address and address prefix for each interface as per [Figure 79](#). (Details not shown.)
2. Configure Router A:
Enable IPv6 multicast routing, enable IPv6 PIM-DM on each interface, and enable MLD on the host-side interface GigabitEthernet 1/0/2.

```
<RouterA> system-view
[RouterA] multicast ipv6 routing-enable
[RouterA] interface gigabitethernet 1/0/1
[RouterA-GigabitEthernet1/0/1] ipv6 pim dm
[RouterA-GigabitEthernet1/0/1] quit
[RouterA] interface gigabitethernet 1/0/2
[RouterA-GigabitEthernet1/0/2] ipv6 pim dm
[RouterA-GigabitEthernet1/0/2] mld enable
```
3. Configure Switch A:
Enable MLD snooping globally.

```
<SwitchA> system-view
[SwitchA] mld-snooping
[SwitchA-mld-snooping] quit
```

Create VLAN 10, assign GigabitEthernet 1/0/1 to VLAN 10, and enable MLD snooping in this VLAN.

```
[SwitchA] vlan 10
[SwitchA-vlan10] port gigabitethernet 1/0/1
[SwitchA-vlan10] mld-snooping enable
```

```
[SwitchA-vlan10] quit
```

Create VLAN 2 and enable MLD snooping in the VLAN.

```
[SwitchA] vlan 2
```

```
[SwitchA-vlan2] mld-snooping enable
```

```
[SwitchA-vlan2] quit
```

The configuration for VLAN 3 and VLAN 4 is similar. (Details not shown.)

Configure GigabitEthernet 1/0/2 as a hybrid port. Configure VLAN 2 as the default VLAN. Configure GigabitEthernet 1/0/2 to permit packets of VLAN 2 to pass and untag the packets when forwarding them.

```
[SwitchA] interface gigabitethernet 1/0/2
```

```
[SwitchA-GigabitEthernet1/0/2] port link-type hybrid
```

```
[SwitchA-GigabitEthernet1/0/2] port hybrid pvid vlan 2
```

```
[SwitchA-GigabitEthernet1/0/2] port hybrid vlan 2 untagged
```

```
[SwitchA-GigabitEthernet1/0/2] port hybrid vlan 10 untagged
```

```
[SwitchA-GigabitEthernet1/0/2] quit
```

The configuration for GigabitEthernet 1/0/3 and GigabitEthernet 1/0/4 is similar. (Details not shown.)

Configure VLAN 10 as an IPv6 multicast VLAN.

```
[SwitchA] multicast-vlan ipv6 10
```

Assign GigabitEthernet 1/0/2 and GigabitEthernet 1/0/3 to IPv6 multicast VLAN 10.

```
[SwitchA-ipv6-mvlan-10] port gigabitethernet 1/0/2 to gigabitethernet 1/0/3
```

```
[SwitchA-ipv6-mvlan-10] quit
```

Assign GigabitEthernet 1/0/4 to IPv6 multicast VLAN 10.

```
[SwitchA] interface gigabitethernet 1/0/4
```

```
[SwitchA-GigabitEthernet1/0/4] port multicast-vlan ipv6 10
```

```
[SwitchA-GigabitEthernet1/0/4] quit
```

4. Verify the configuration:

View the IPv6 multicast VLAN information on Switch A.

```
[SwitchA] display multicast-vlan ipv6
```

```
Total 1 IPv6 multicast-vlan(s)
```

```
IPv6 Multicast vlan 10
```

```
  subvlan list:
```

```
    no subvlan
```

```
  port list:
```

```
    GE1/0/2
```

```
    GE1/0/3
```

```
    GE1/0/4
```

View the MLD snooping multicast group information on Switch A.

```
[SwitchA] display mld-snooping group
```

```
Total 1 IP Group(s).
```

```
Total 1 IP Source(s).
```

```
Total 1 MAC Group(s).
```

```
Port flags: D-Dynamic port, S-Static port, C-Copy port, P-PIM port
```

```
Subvlan flags: R-Real VLAN, C-Copy VLAN
```

```
Vlan(id):10.
```

```
  Total 1 IP Group(s).
```

```
  Total 1 IP Source(s).
```

```
Total 1 MAC Group(s).
Router port(s):total 1 port(s).
    GE1/0/1                (D)
IP group(s):the following ip group(s) match to one mac group.
  IP group address:FF1E::101
    (::, FF1E::101):
      Host port(s):total 3 port(s).
        GE1/0/2            (D)
        GE1/0/3            (D)
        GE1/0/4            (D)
MAC group(s):
  MAC group address:3333-0000-0101
    Host port(s):total 3 port(s).
      GE1/0/2
      GE1/0/3
      GE1/0/4
```

The output shows that MLD snooping is maintaining router ports and member ports in VLAN 10.

Configuring IPv6 multicast routing and forwarding (available only on the HP 5500 EI)

Overview

In IPv6 multicast implementations, the following types of tables implement multicast routing and forwarding:

- **Multicast routing table of an IPv6 multicast routing protocol**—Each IPv6 multicast routing protocol has its own multicast routing table, such as IPv6 PIM routing table.
- **General IPv6 multicast routing table**—The multicast routing information of different IPv6 multicast routing protocols forms a general IPv6 multicast routing table.
- **IPv6 multicast forwarding table**—The IPv6 multicast forwarding table guides the forwarding of IPv6 multicast packets.

An IPv6 multicast forwarding table consists of a set of (S, G) entries. Each entry indicates the routing information for delivering multicast data from a multicast source to a multicast group. If a router supports multiple IPv6 multicast protocols, its IPv6 multicast routing table will include routes that these protocols have generated. The router chooses the optimal route from the IPv6 multicast routing table based on the configured multicast routing and forwarding policy and installs the route entry into its IPv6 multicast forwarding table.

The term "router" in this document refers to both routers and Layer 3 switches.

The term "interface" in the IPv6 multicast routing and forwarding features refers to Layer 3 interfaces, including VLAN interfaces and route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).

RPF check mechanism

An IPv6 multicast routing protocol relies on the existing IPv6 unicast routing information or IPv6 MBGP routes in creating IPv6 multicast routing entries. When creating IPv6 multicast routing table entries, an IPv6 multicast routing protocol uses the reverse path forwarding (RPF) check mechanism to ensure IPv6 multicast data delivery along the correct path. The RPF check mechanism also helps avoid data loops caused by various reasons.

RPF check process

An RPF check is based on one of the following routing tables:

- **IPv6 unicast routing table**—Contains the shortest path to each destination subnet
- **IPv6 MBGP routing table**—Contains IPv6 multicast routing information

When a router performs an RPF check, it searches its IPv6 unicast routing table and IPv6 MBGP routing table at the same time. The specific process is as follows:

1. The router chooses each optimal route from the IPv6 unicast routing table and the IPv6 MBGP routing table:

- The router searches its IPv6 unicast routing table by using the IPv6 address of the packet source as the destination address and automatically selects the optimal route as the RPF route. The outgoing interface in the corresponding routing entry is the RPF interface and the next hop is the RPF neighbor. The router considers the path along which the IPv6 multicast packet from the RPF neighbor arrived on the RPF interface to be the shortest path that leads back to the source.
 - The router automatically chooses an optimal IPv6 MBGP route by searching its MBGP routing table, and using the IPv6 address of the packet source as the destination address. The outgoing interface in the corresponding routing entry is the RPF interface and the next hop is the RPF neighbor.
2. The router selects one of these optimal routes as the RPF route. The selection process is as follows:
 - If configured to use the longest match principle, the router selects the longest match route from these optimal routes. If these routes have the same prefix length, the router selects the route with a higher priority. If these routes have the same priority, the router selects the IPv6 MBGP route as the RPF route.
 - If not configured to use the longest match principle, the router selects the route with a higher priority. If these routes have the same priority, the router selects the IPv6 MBGP route as the RPF route.

The term "packet source" can mean different things in different situations:

- For a packet that traveling along the shortest path tree (SPT) from the multicast source to the receivers or the rendezvous point (RP), the packet source for RPF check is the multicast source.
- For a packet that traveling along the rendezvous point tree (RPT) from the RP to the receivers, or along the source-side RPT from the multicast source to the RP, the packet source for RPF check is the RP.
- For a bootstrap message from the bootstrap router (BSR), the packet source for RPF check is the BSR.

For more information about the concepts of SPT, RPT, source-side RPT, RP, and BSR, see "[Configuring IPv6 PIM \(available only on the HP 5500 EI\)](#)."

RPF check implementation in IPv6 multicast

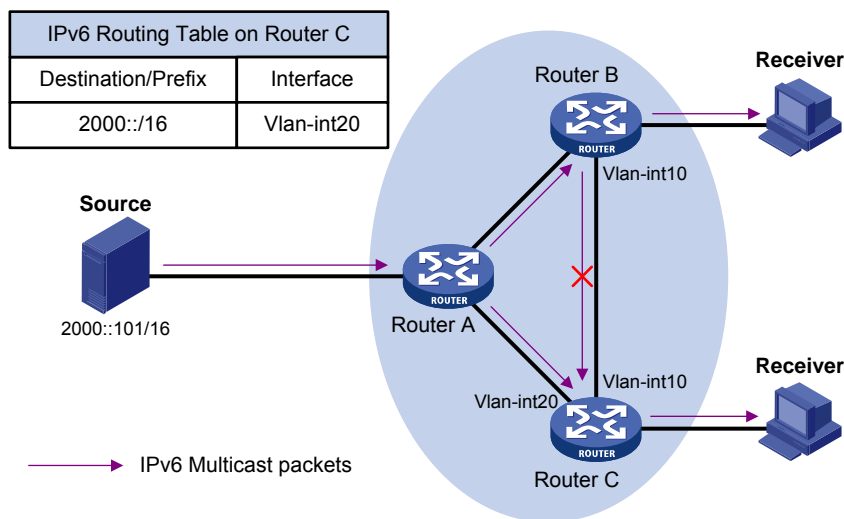
Implementing an RPF check on each received IPv6 multicast data packet would heavily burden the router. The use of an IPv6 multicast forwarding table is the solution to this issue. When creating an IPv6 multicast routing entry and an IPv6 multicast forwarding entry for an IPv6 multicast packet, the router sets the RPF interface of the packet as the incoming interface of the (S, G) entry. After receiving an (S, G) IPv6 multicast packet, the router first searches its IPv6 multicast forwarding table:

1. If the corresponding (S, G) entry does not exist in the IPv6 multicast forwarding table, the packet undergoes an RPF check. The router creates an IPv6 multicast routing entry based on the relevant routing information and installs the entry into the IPv6 multicast forwarding table, with the RPF interface as the incoming interface.
 - If the interface that received the packet is the RPF interface, the RPF check succeeds and the router forwards the packet to all the outgoing interfaces.
 - If the interface that received the packet is not the RPF interface, the RPF check fails and the router discards the packet.
2. If the corresponding (S, G) entry exists, and the interface that received the packet is the incoming interface, the router forwards the packet to all the outgoing interfaces.
3. If the corresponding (S, G) entry exists, but the interface that received the packet is not the incoming interface in the IPv6 multicast forwarding table, the IPv6 multicast packet undergoes an RPF check.

- If the RPF interface is the incoming interface of the (S, G) entry, this means the (S, G) entry is correct but the packet arrived from a wrong path. The packet will be discarded.
- If the RPF interface is not the incoming interface, this means that the (S, G) entry has expired, and the router replaces the incoming interface with the RPF interface. If the interface on which the packet arrived is the RPF interface, the router forwards the packet to all the outgoing interfaces. Otherwise it discards the packet.

Assume that IPv6 unicast routes are available in the network, IPv6 MBGP is not configured, and IPv6 multicast packets travel along the SPT from the multicast source to the receivers, as shown in Figure 80. The IPv6 multicast forwarding table on Router C contains the (S, G) entry, with VLAN-interface 20 as the RPF interface.

Figure 80 RPF check process



- When an IPv6 multicast packet arrives on VLAN-interface 20 of Router C, because the interface is the incoming interface of the (S, G) entry, the router forwards the packet to all outgoing interfaces.
- When an IPv6 multicast packet arrives on VLAN-interface 10 of Router C, because the interface is not the incoming interface of the (S, G) entry, the router performs an RPF check on the packet. The router searches its IPv6 unicast routing table and finds that the outgoing interface to Source (the RPF interface) is VLAN-interface 20. This means that the (S, G) entry is correct and the packet arrived along a wrong path. The RPF check fails and the packet is discarded.

Configuration task list

Task	Remarks	
Enabling IPv6 multicast routing	Required	
Configuring IPv6 multicast routing and forwarding	Configuring an IPv6 multicast routing policy	Optional
	Configuring an IPv6 multicast forwarding range	Optional
	Configuring the IPv6 multicast forwarding table size	Optional

Enabling IPv6 multicast routing

Before you configure any Layer 3 IPv6 multicast functionality, you must enable IPv6 multicast routing.

To enable IPv6 multicast routing:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IPv6 multicast routing.	multicast ipv6 routing-enable	Disabled by default.

Configuring IPv6 multicast routing and forwarding

Configuration prerequisites

Before you configure IPv6 multicast routing and forwarding, complete the following tasks:

- Enable IPv6 forwarding and configure an IPv6 unicast routing protocol so that all devices in the domain are interoperable at the network layer.
- Configure IPv6 PIM-DM or IPv6 PIM-SM.
- Determine the maximum number of downstream nodes for a single entry in the IPv6 multicast forwarding table.
- Determine the maximum number of entries in the IPv6 multicast forwarding table.

Configuring an IPv6 multicast routing policy

You can configure the router to determine the RPF route based on the longest match principle. For more information about RPF route selection, see "[RPF check process](#)."

By configuring per-source or per-source-and-group load splitting, you can optimize the traffic delivery when multiple IPv6 multicast data streams are handled.

To configure an IPv6 multicast routing policy:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the device to select the RPF route based on the longest match.	multicast ipv6 longest-match	Optional. The route with the highest priority is selected as the RPF route by default.
3. Configure IPv6 multicast load splitting.	multicast ipv6 load-splitting { source source-group }	Optional. Disabled by default.

Configuring an IPv6 multicast forwarding range

IPv6 multicast packets do not travel infinitely in a network. The IPv6 multicast data of each IPv6 multicast group must be transmitted within a definite scope.

You can configure the forwarding boundary for a specific IPv6 multicast group or an IPv6 multicast group with the scope field in its group address being specified on all interfaces that support IPv6 multicast forwarding. A multicast forwarding boundary sets the boundary condition for the IPv6 multicast groups in the specified range or scope. If the destination address of an IPv6 multicast packet matches the set boundary condition, the packet will not be forwarded. Once an IPv6 multicast boundary is configured on an interface, this interface can no longer forward IPv6 multicast packets (including those sent from the local device) or receive IPv6 multicast packets.

To configure an IPv6 multicast forwarding range:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure an IPv6 multicast forwarding boundary.	multicast ipv6 boundary { <i>ipv6-group-address prefix-length</i> scope { <i>scope-id</i> admin-local global organization-local site-local } }	No forwarding boundary by default.

Configuring the IPv6 multicast forwarding table size

The switch maintains the corresponding forwarding entry for each IPv6 multicast packet that it receives. Excessive IPv6 multicast routing entries, however, can exhaust the switch's memory and cause lower performance. You can set a limit on the number of entries in the IPv6 multicast forwarding table based on the actual networking situation and the performance requirements. If the configured maximum number of IPv6 multicast forwarding table entries is smaller than the current value, the entries in excess are not immediately deleted. Instead, the IPv6 multicast routing protocol that runs on the switch deletes them. The switch no longer adds new IPv6 multicast forwarding entries until the number of existing IPv6 multicast forwarding entries comes down below the configured value.

When the switch forwards IPv6 multicast traffic, it replicates a copy of the IPv6 multicast traffic for each downstream node and forwards the traffic. Each of these downstream nodes forms a branch of the IPv6 multicast distribution tree. You can configure the maximum number of downstream nodes (the maximum number of outgoing interfaces) for a single entry in the IPv6 multicast forwarding table to lessen the burden on the switch for replicating IPv6 multicast traffic. If the configured maximum number of downstream nodes for a single IPv6 multicast forwarding entry is smaller than the current number, the downstream nodes in excess are not deleted immediately. Instead, the IPv6 multicast routing protocol deletes them. The switch no longer adds new IPv6 multicast forwarding entries for newly added downstream nodes until the number of existing downstream nodes comes down below the configured value.

To configure the IPv6 multicast forwarding table size:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the maximum number of entries in the IPv6 multicast forwarding table.	multicast ipv6 forwarding-table route-limit <i>limit</i>	Optional. 1000 by default.

Step	Command	Remarks
3. Configure the maximum number of downstream nodes for a single IPv6 multicast forwarding entry.	multicast ipv6 forwarding-table downstream-limit <i>limit</i>	Optional. 128 by default.

Displaying and maintaining IPv6 multicast routing and forwarding

Task	Command	Remarks
Display the IPv6 multicast boundary information.	display multicast ipv6 boundary { group [<i>ipv6-group-address</i> [<i>prefix-length</i>]] scope [<i>scope-id</i>] } [interface <i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the information of the IPv6 multicast forwarding table.	display multicast ipv6 forwarding-table [<i>ipv6-source-address</i> [<i>prefix-length</i>] <i>ipv6-group-address</i> [<i>prefix-length</i>] incoming-interface { <i>interface-type interface-number</i> register } outgoing-interface { exclude include match } { <i>interface-type interface-number</i> register } statistics slot <i>slot-number</i>] * [port-info] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the DF information of the IPv6 multicast forwarding table.	display multicast ipv6 forwarding-table df-info [<i>rp-address</i>] [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the information of the IPv6 multicast routing table.	display multicast ipv6 routing-table [<i>ipv6-source-address</i> [<i>prefix-length</i>] <i>ipv6-group-address</i> [<i>prefix-length</i>] incoming-interface { <i>interface-type interface-number</i> register } outgoing-interface { exclude include match } { <i>interface-type interface-number</i> register }] * [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the RPF route information of the specified IPv6 multicast source.	display multicast ipv6 rpf-info <i>ipv6-source-address</i> [<i>ipv6-group-address</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear forwarding entries from the IPv6 multicast forwarding table.	reset multicast ipv6 forwarding-table { { <i>ipv6-source-address</i> [<i>prefix-length</i>] <i>ipv6-group-address</i> [<i>prefix-length</i>] incoming-interface { <i>interface-type interface-number</i> register } } * all }	Available in user view

Task	Command	Remarks
Clear routing entries from the IPv6 multicast routing table.	reset multicast ipv6 routing-table { { <i>ipv6-source-address</i> [<i>prefix-length</i>] <i>ipv6-group-address</i> [<i>prefix-length</i>] incoming-interface { <i>interface-type</i> <i>interface-number</i> register } } * all }	Available in user view

For more information about designated forwarder (DF), see "[Configuring IPv6 PIM \(available only on the HP 5500 EI\)](#)."

ⓘ IMPORTANT:

The **reset** command clears the information in the IPv6 multicast routing table or the multicast forwarding table and might cause transmission failure of IPv6 multicast information.

When a routing entry is deleted from the IPv6 multicast routing table, the corresponding forwarding entry is also deleted from the IPv6 multicast forwarding table.

When a forwarding entry is deleted from the IPv6 multicast forwarding table, the corresponding routing entry is also deleted from the IPv6 multicast routing table.

Troubleshooting IPv6 multicast policy configuration

Abnormal termination of IPv6 multicast data

Symptom

- A host sends an MLD report announcing its joining an IPv6 multicast group (G). However, no member information about the IPv6 multicast group (G) exists on the intermediate router. The intermediate router can receive IPv6 multicast packets successfully, but the packets cannot reach the stub network.
- The interface of the intermediate router receives the IPv6 multicast packets, but no corresponding (S, G) entry exists in the IPv6 PIM routing table.

Analysis

- The **multicast ipv6 boundary** command filters IPv6 multicast packets received on an interface. If an IPv6 multicast packet fails to match the IPv6 ACL rule of this command, IPv6 PIM will create no routing entry.
- In addition, the **source-policy** command in IPv6 PIM filters received IPv6 multicast packets. If an IPv6 multicast packet fails to match the IPv6 ACL rule of this command, IPv6 PIM will not create a routing entry, either.

Solution

1. Use the **display current-configuration** command to display the IPv6 ACL rule configured on the multicast forwarding boundary. Change the IPv6 ACL rule used in the **multicast ipv6 boundary** command so that the source IP address of the IPv6 multicast packets and the IPv6 multicast group address can both match the IPv6 ACL rule.
2. View the configuration of the multicast filter. Use the **display current-configuration** command to display the configuration of the IPv6 multicast filter. Change the IPv6 ACL rule used in the

source-policy command so that the source IP address of the IPv6 multicast packets and the IPv6 multicast group address can both match the IPv6 ACL rule.

Configuring MLD (available only on the HP 5500 EI)

Overview

An IPv6 router uses the Multicast Listener Discovery (MLD) protocol to discover the presence of multicast listeners on the directly attached subnets. Multicast listeners are nodes that want to receive IPv6 multicast packets.

Through MLD, the router can determine whether any IPv6 multicast listeners exist on the directly connected subnets, put corresponding records in the database, and maintain timers related to IPv6 multicast addresses.

Routers running MLD use an IPv6 unicast link-local address as the source address to send MLD messages. MLD messages are Internet Control Message Protocol for IPv6 (ICMPv6) messages. All MLD messages are confined to the local subnet, with a hop count of 1.

The term "router" in this document refers to both routers and Layer 3 switches.

The term "interface" in the MLD features refers to Layer 3 interfaces, including VLAN interfaces and route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).

MLD versions

- MLDv1 (defined in RFC 2710), which is derived from IGMPv2.
- MLDv2 (defined in RFC 3810), which is derived from IGMPv3.

All MLD versions support the Any-Source Multicast (ASM) model. In addition, MLDv2 can directly implement the Source-Specific Multicast (SSM) model, but MLDv1 must work with the MLD SSM mapping function to implement SSM service.

For more information about the ASM and SSM models, see "[Multicast overview](#)."

How MLDv1 works

MLDv1 implements IPv6 multicast listener management based on the query/response mechanism.

MLD querier election

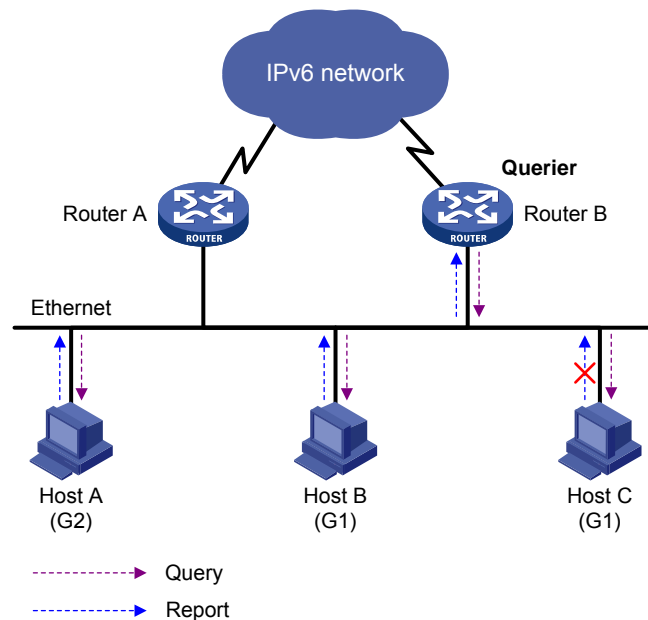
All IPv6 multicast routers on the same subnet can monitor MLD listener report messages (often called "reports") from hosts, but only one router is needed for sending MLD query messages (often called "queries"). A querier election mechanism determines which router will act as the MLD querier on the subnet.

1. Initially, every MLD router assumes itself as the querier and sends MLD general query messages (often called "general queries") to all hosts and routers on the local subnet. The destination address is FF02::1.

2. After receiving a general query, every MLD router compares the source IPv6 address of the query message with its own interface address. After comparison, the router with the lowest IPv6 address wins the querier election and all other routers become non-queriers.
3. All the non-queriers start a timer called the "other querier present timer." If a router receives an MLD query from the querier before the timer expires, it resets this timer. Otherwise, it assumes that the querier has timed out and initiates a new querier election process.

Joining an IPv6 multicast group

Figure 81 MLD queries and reports



Assume that Host B and Host C will receive IPv6 multicast data addressed to IPv6 multicast group G1, and Host A will receive IPv6 multicast data addressed to G2, as shown in Figure 81. The following process describes how the hosts join the IPv6 multicast groups and how the MLD querier (Router B in the figure) maintains the IPv6 multicast group memberships:

1. The hosts send unsolicited MLD reports to the addresses of the IPv6 multicast groups that they will join, without having to wait for the MLD queries from the MLD querier.
2. The MLD querier periodically multicasts MLD queries (with the destination address of FF02::1) to all hosts and routers on the local subnet.
3. After receiving a query message, Host B or Host C (the delay timer of whichever expires first) sends an MLD report to the IPv6 multicast group address of G1, to announce its membership for G1. Assume that Host B sends the report message. After hearing the report from Host B, Host C, which is on the same subnet as Host B, suppresses its own report for G1, because the MLD routers (Router A and Router B) have already known that at least one host on the local subnet is interested in G1. This mechanism, known as the "MLD report suppression", helps reduce traffic on the local subnet.
4. At the same time, because Host A is interested in G2, it sends a report to the IPv6 multicast group address of G2.
5. Through the query/report process, the MLD routers learn that members of G1 and G2 are attached to the local subnet, and the IPv6 multicast routing protocol (for example, IPv6 PIM) that is running on the routers generates (*, G1) and (*, G2) multicast forwarding entries. These entries will be the basis for subsequent IPv6 multicast forwarding, where * represents any IPv6 multicast source.

6. When the IPv6 multicast data addressed to G1 or G2 reaches an MLD router, because the (*, G1) and (*, G2) multicast forwarding entries exist on the MLD router, the router forwards the IPv6 multicast data to the local subnet, and then the receivers on the subnet receive the data.

Leaving an IPv6 multicast group

When a host leaves a multicast group, the following occur:

1. The host sends an MLD done message to all IPv6 multicast routers on the local subnet. The destination address is FF02::2.
2. After receiving the MLD done message, the querier sends a configurable number of multicast-address-specific queries to the group that the host is leaving. The destination address field and group address field of the message are both filled with the address of the IPv6 multicast group that is being queried.
3. One of the remaining members (if any on the subnet) of the group being queried should send a report within the time of the maximum response delay set in the query messages.
4. If the querier receives a report for the group within the maximum response delay time, it will maintain the memberships of the IPv6 multicast group. Otherwise, the querier will assume that no hosts on the subnet are still interested in IPv6 multicast traffic addressed to that group and will stop maintaining the memberships of the group.

How MLDv2 works

Compared with MLDv1, MLDv2 provides the following new features:

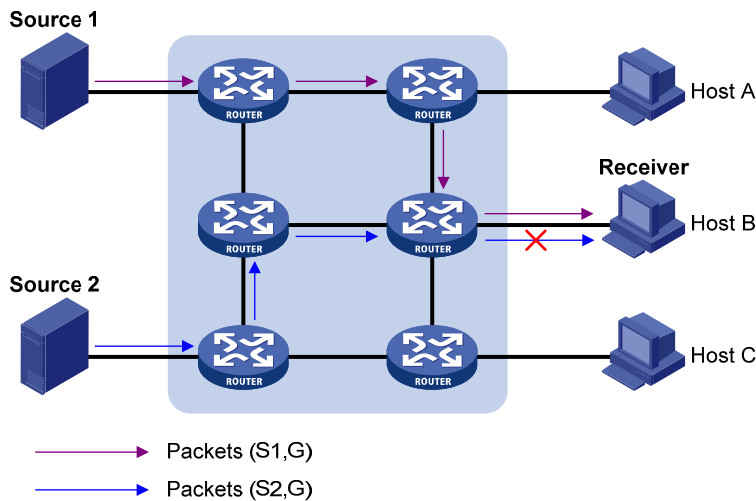
IPv6 multicast group filtering

MLDv2 has introduced IPv6 multicast source filtering modes (Include and Exclude), so that a host not only can join a designated IPv6 multicast group, but also can specify to receive or reject multicast data from designated IPv6 multicast sources. When a host joins an IPv6 multicast group, one of the following situation occurs:

- If it expects IPv6 multicast data from specific IPv6 multicast sources like S1, S2, ..., it sends a report with Filter-Mode denoted as "Include Sources (S1, S2, ...)."
- If it does not expect IPv6 multicast data from specific IPv6 multicast sources like S1, S2, ..., it sends a report with Filter-Mode denoted as "Exclude Sources (S1, S2, ...)."

As shown in [Figure 82](#), the network comprises two IPv6 multicast sources, Source 1 (S1) and Source 2 (S2), both of which can send IPv6 multicast data to IPv6 multicast group G. Host B is interested only in the IPv6 multicast data that Source 1 sends to G but not in the data from Source 2.

Figure 82 Flow paths of multicast-address-and-source-specific multicast traffic



In the case of MLDv1, Host B cannot select IPv6 multicast sources when it joins IPv6 multicast group G. Therefore, IPv6 multicast streams from both Source 1 and Source 2 will flow to Host B whether it needs them or not.

When MLDv2 is running on the hosts and routers, Host B can explicitly express its interest in the IPv6 multicast data that Source 1 sends to G (denoted as (S1, G)), rather than the IPv6 multicast data that Source 2 sends to G (denoted as (S2, G)). Thus, only IPv6 multicast data from Source 1 will be delivered to Host B.

MLD state

A multicast router that is running MLDv2 maintains the multicast address state per multicast address per attached subnet. The multicast address state consists of the following information:

- **Filter mode**—The router keeps tracing the Include or Exclude state.
- **List of sources**—The router keeps tracing the newly added or deleted IPv6 multicast source.
- **Timers**—Filter timers, including the time that the router waits before switching to the Include mode after an IPv6 multicast address times out, the source timer for source recording, and so on.

Receiver host state listening

By listening to the state of receiver hosts, a multicast router running MLDv2 records and maintains information of hosts joining the source group on the attached subnet.

MLD messages

The following descriptions are based on MLDv2 messages.

MLD query message

An MLD querier learns the multicast listening state of neighbor interfaces by sending MLD query messages. The dark area in Figure 83 shows the MLDv1 message format.

Figure 83 MLDv2 query message format

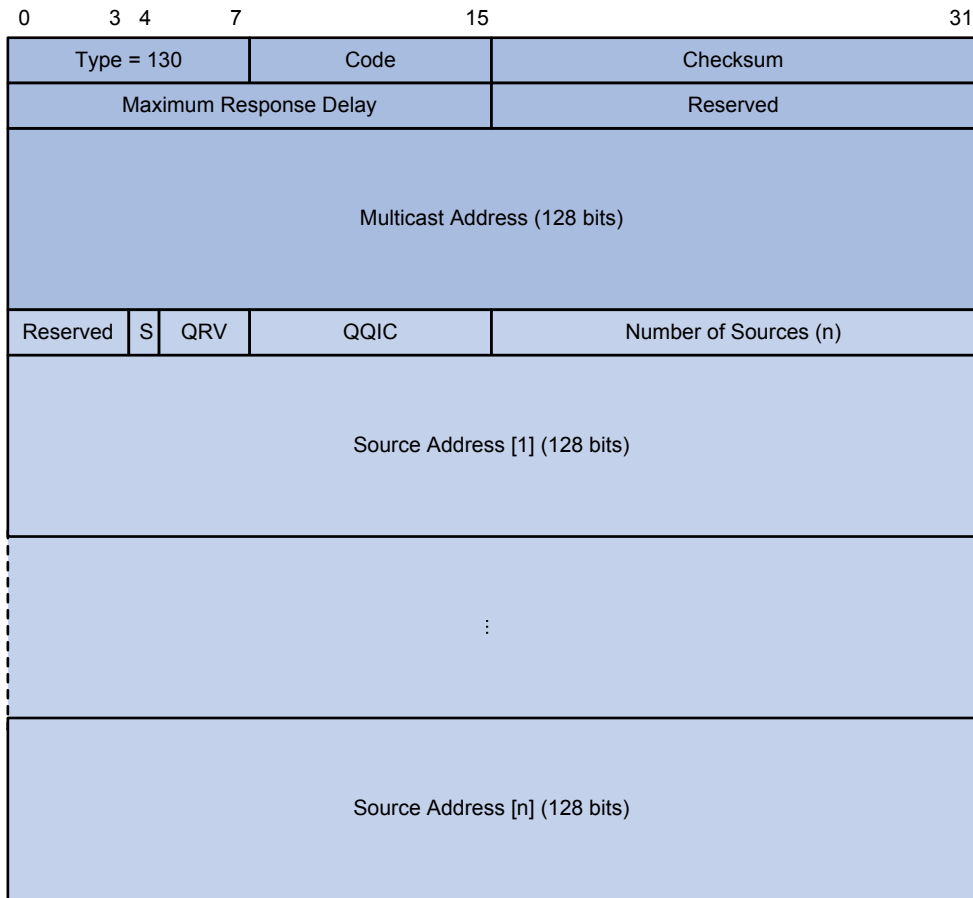


Table 9 MLDv2 query message field description

Field	Description
Type = 130	Message type. For a query message, this field is set to 130.
Code	Initialized to zero.
Checksum	Standard IPv6 checksum.
Maximum Response Delay	Maximum response delay allowed before a host sends a report message.
Reserved	Reserved field and initialized to zero.
Multicast Address	<ul style="list-style-type: none"> This field is set to 0 in a general query message. It is set to a specific IPv6 multicast address in a multicast-address-specific query message or multicast-address-and-source-specific query message.
S	Flag indicating whether a router updates the timer for suppression after receiving a query message.
QRV	Querier's Robustness Variable.
QQIC	Querier's Query Interval Code.

Field	Description
Number of Sources	<ul style="list-style-type: none"> This field is set to 0 in a general query message or a multicast-address-specific query message. This field represents the number of source addresses in a multicast-address-and-source-specific query message.
Source Address(i)	IPv6 multicast source address in a multicast-address-specific query message (i = 1, 2, ..., n, where n represents the number of multicast source addresses.)

MLD report message

A host sends an MLD report message to report the current multicast listening state. Figure 84 shows the MLD report message format.

Figure 84 MLDv2 report message format

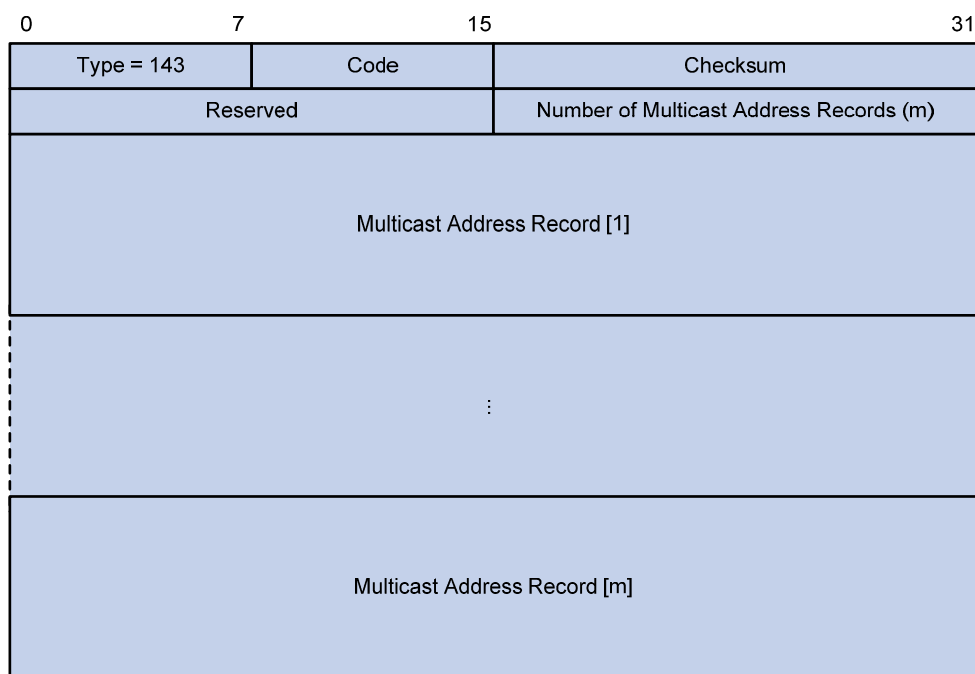


Table 10 MLDv2 report message field description

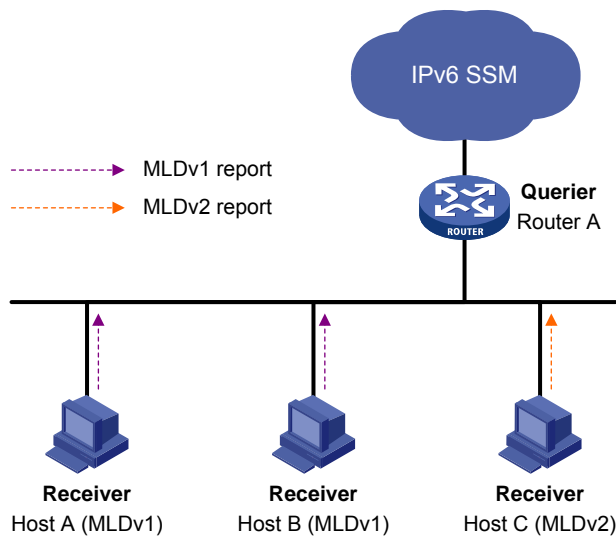
Field	Description
Type = 143	Message type. For a report message, this field is set to 143.
Reserved	The Reserved fields are set to 0 on transmission and ignored on reception.
Checksum	Standard IPv6 checksum.
Number of Multicast Address Records	This field indicates how many IPv6 multicast address records are present in this report message.
Multicast Address Record(i)	This field represents information of each IPv6 multicast address the host listens to on the interface from which the report message is sent, including record type, IPv6 multicast address, and IPv6 multicast source address on the sender (i= 1, 2, ... m, where m represents the number of IPv6 multicast address records).

MLD SSM mapping

The MLD SSM mapping feature enables you to configure static MLD SSM mappings on the last hop router to provide SSM support for receiver hosts that are running MLDv1. The SSM model assumes that the last hop router has identified the desired IPv6 multicast sources when receivers join IPv6 multicast groups.

- When an MLDv2 enabled host joins a multicast group, it can explicitly specify one or more multicast sources in its MLDv2 report.
- An MLDv1-enabled host, however, cannot specify multicast source addresses in its MLDv1 report. You must configure the MLD SSM mapping feature to translate the (*, G) information in the MLDv1 report into (G, INCLUDE, (S1, S2...)) information.

Figure 85 Network diagram



On the IPv6 SSM network in [Figure 85](#), Host A and Host B are running MLDv1 and Host C is running MLDv2. To provide SSM service for Host A and Host B, you must configure the MLD SSM mapping feature on Router A.

With the MLD SSM mapping feature configured, when Router A receives an MLDv1 report, it checks the IPv6 multicast group address G carried in the message.

- If G is not in the IPv6 SSM group range, Router A cannot provide the SSM service but can provide the ASM service.
- If G is in the IPv6 SSM group range but no MLD SSM mappings have been configured for the IPv6 multicast group G on Router A, Router A cannot provide SSM service and drops the packet.
- If G is in the IPv6 SSM group range, and the MLD SSM mappings have been configured on Router A for multicast group G, Router A translates the (*, G) information in the MLD report into (G, INCLUDE, (S1, S2...)) information based on the configured MLD SSM mappings and provides SSM service accordingly.

NOTE:

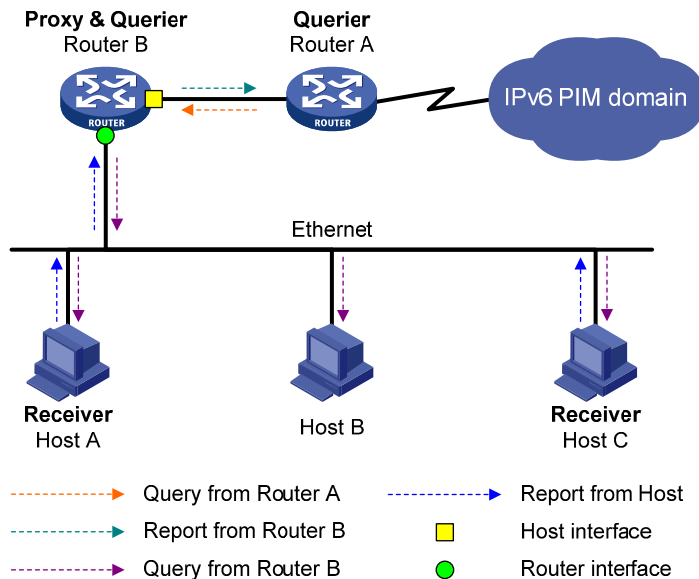
The MLD SSM mapping feature does not process MLDv2 reports.

For more information about the IPv6 SSM group range, see "[Configuring IPv6 PIM \(available only on the HP 5500 EI\)](#)."

MLD proxying

In some simple tree-shaped topologies, you do not need to configure complex IPv6 multicast routing protocols, such as IPv6 PIM, on the boundary devices. Instead, you can configure MLD proxying on these devices. With MLD proxying configured, the device serves as a proxy for the downstream hosts to send MLD messages, maintain group memberships, and implement IPv6 multicast forwarding based on the memberships. In this case, the MLD proxy device is a host but no longer an IPv6 PIM neighbor to the upstream device.

Figure 86 Network diagram



As shown in [Figure 86](#), an MLD proxy device has the following types of interfaces:

- **Upstream interface**—Also called the "proxy interface." A proxy interface is an interface on which MLD proxying is configured. It is in the direction toward the root of the multicast forwarding tree. An upstream interface acts as a host that is running MLD, and is also called a "host interface."
- **Downstream interface**—An interface that is running MLD and not in the direction toward the root of the multicast forwarding tree. A downstream interface acts as a router that is running MLD, and is also called a "router interface."

A device with MLD proxying configured maintains a group membership database, which stores the group memberships on all the downstream interfaces in this database. Each entry comprises the multicast address, filter mode, and source list. Such an entry is a collection of members in the same multicast group on each downstream interface.

A proxy device performs host functions on the upstream interface based on the database. It responds to the queries according to the information in the database or sends join/leave messages when the database changes. The proxy device performs router functions on the downstream interfaces by participating in the querier election, sending queries, and maintaining memberships based on the reports.

Protocols and standards

- RFC 2710, *Multicast Listener Discovery (MLD) for IPv6*
- RFC 3810, *Multicast Listener Discovery Version 2 (MLDv2) for IPv6*

- RFC 4605, *Internet Group Management Protocol (IGMP)/Multicast Listener Discovery (MLD)-Based Multicast Forwarding ("IGMP/MLD Proxying")*

MLD configuration task list

Task	Remarks	
	Enabling MLD	Required
Configuring basic MLD functions	Configuring the MLD version	Option
	Configuring static joining	Optional
	Configuring an IPv6 multicast group filter	Optional
	Setting the maximum number of IPv6 multicast groups that an interface can join	Optional
	Configuring MLD message options	Optional
Adjusting MLD performance	Configuring MLD query and response parameters	Optional
	Configuring MLD fast-leave processing	Optional
	Enabling the MLD host tracking function	Optional
	Setting the DSCP value for MLD messages	Optional
Configuring MLD SSM mapping	Enabling MLD SSM mapping	Optional
	Configuring MLD SSM mappings	Optional
Configuring MLD proxying	Enabling MLD proxying	Optional
	Configuring IPv6 multicast forwarding on a downstream interface	Optional

NOTE:

- In MLD view, the configuration is effective globally. In interface view, the configuration is effective on only the current interface.
- If no configuration is performed in interface view, the global configuration in MLD view will apply to that interface. Configurations performed in interface view take precedence over those performed in MLD view.

Configuring basic MLD functions

Configuration prerequisites

Before you configure basic MLD functions, complete the following tasks:

- Enable IPv6 forwarding and configure an IPv6 unicast routing protocol so that all devices in the domain can be interoperable at the network layer.
- Configure IPv6 PIM-DM or IPv6 PIM-SM.
- Determine the MLD version.

- Determine the IPv6 multicast group address and IPv6 multicast source address for static group member configuration.
- Determine the ACL rule for IPv6 multicast group filtering.
- Determine the maximum number of IPv6 multicast groups that an interface can join.

Enabling MLD

Enable MLD on the interface on which IPv6 multicast group memberships will be created and maintained.

To enable MLD:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IPv6 multicast routing.	multicast ipv6 routing-enable	Disable by default
3. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Enable MLD.	mld enable	Disabled by default

For more information about the **multicast ipv6 routing-enable** command, see *IP Multicast Command Reference*.

Configuring the MLD version

Because MLD message types and formats vary with MLD versions, the same MLD version should be configured for all routers on the same subnet before MLD can work properly.

Configuring an MLD version globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter MLD view.	mld	N/A
3. Configure an MLD version globally.	version <i>version-number</i>	MLDv1 by default

Configuring an MLD version on an interface

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure an MLD version on the interface.	mld version <i>version-number</i>	MLDv1 by default

Configuring static joining

After an interface is configured as a static member of an IPv6 multicast group or an IPv6 multicast source and group, it will act as a virtual member of the IPv6 multicast group to receive IPv6 multicast data addressed to that IPv6 multicast group for the purpose of testing IPv6 multicast data forwarding.

Configuration guidelines

Before you can configure an interface of an IPv6 PIM-SM device as a static member of an IPv6 multicast group or an IPv6 multicast source and group, if the interface is IPv6 PIM-SM enabled, it must be an IPv6 PIM-SM DR. If this interface is MLD enabled but not IPv6 PIM-SM enabled, it must be an MLD querier. For more information about IPv6 PIM-SM and a DR, see "[Configuring IPv6 PIM \(available only on the HP 5500 EI\)](#)."

A static member port does not respond to queries from the MLD querier. When you configure an interface as a static member port or remove this configuration on the interface, the interface does not unsolicitedly send any MLD report or an MLD done message. In other words, the interface is not a real member of the IPv6 multicast group or the IPv6 multicast and source group.

Configuration procedure

To configure a static member of an IPv6 multicast group or an IPv6 multicast source and group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure a static member of an IPv6 multicast group or an IPv6 multicast source and group.	mld static-group <i>ipv6-group-address</i> [source <i>ipv6-source-address</i>]	By default, an interface is not a static member of any IPv6 multicast group or IPv6 multicast source and group.

Configuring an IPv6 multicast group filter

To restrict the hosts on the network attached to an interface from joining certain IPv6 multicast groups, you can set an IPv6 ACL rule on the interface so that the interface maintains only the IPv6 multicast groups matching the criteria.

To configure an IPv6 multicast group filter:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure an IPv6 multicast group filter.	mld group-policy <i>acl6-number</i> [<i>version-number</i>]	By default, no IPv6 group filter is configured on the interface. That is, hosts on the current interface can join any valid multicast group.

Setting the maximum number of IPv6 multicast groups that an interface can join

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the maximum number of IPv6 multicast groups that the interface can join.	mld group-limit <i>limit</i>	1000 by default.

NOTE:

This configuration takes effect for dynamically joined IPv6 multicast groups but not the statically configured multicast groups.

Adjusting MLD performance

For the configuration tasks in this section:

- In MLD view, the configuration is effective globally. In interface view, the configuration is effective only on the current interface.
- If the same function or parameter is configured in both MLD view and interface view, the configuration performed in interface view is given priority, regardless of the configuration order.

Configuration prerequisites

Before adjusting MLD performance, complete the following tasks:

- Enable IPv6 forwarding and configure an IPv6 unicast routing protocol so that all devices in the domain can be interoperable at the network layer.
- Configure basic MLD functions.
- Determine the startup query interval.
- Determine the startup query count.
- Determine the MLD query interval.
- Determine the MLD querier's robustness variable.
- Determine the maximum response delay of MLD general query messages.
- Determine the MLD last listener query interval.
- Determine the MLD other querier present interval.
- Determine the DSCP value for MLD messages.

Configuring MLD message options

MLD queries include multicast-address-specific queries and multicast-address-and-source-specific queries, and IPv6 multicast groups change dynamically, so a device cannot maintain the information for all IPv6 multicast sources and groups. Therefore, a router might receive IPv6 multicast packets addressed to IPv6

multicast groups that have no members on the local subnet. In this case, the Router-Alert option carried in the IPv6 multicast packets is useful for the router to determine whether to deliver the IPv6 multicast packets to the upper-layer protocol for processing. For more information about the Router-Alert option, see RFC 2113.

An MLD message is processed differently depending on whether it carries the Router-Alert option in the IPv6 header, as follows:

- For compatibility, the device by default ignores the Router-Alert option and processes all received MLD messages, no matter whether the MLD messages carry the Router-Alert option or not.
- To enhance device performance, avoid unnecessary costs, and ensure protocol security, configure the device to discard MLD messages that do not carry the Router-Alert option.

Configuring the Router-Alert option for MLD messages globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter MLD view.	mld	N/A
3. Configure the interface to discard any MLD message without the Router-Alert option.	require-router-alert	By default, the device does not check MLD messages for the Router-Alert option.
4. Enable the insertion of the Router-Alert option into MLD messages.	send-router-alert	By default, MLD messages carry the Router-Alert option.

Configuring the Router-Alert option on an interface

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the interface to discard any MLD message without the Router-Alert option.	mld require-router-alert	By default, the device does not check MLD messages for the Router-Alert option.
4. Enable the insertion of the Router-Alert option into MLD messages.	mld send-router-alert	By default, MLD messages carry the Router-Alert option.

Configuring MLD query and response parameters

MLD query and response parameters

On startup, the MLD querier sends MLD general queries at the startup query interval, which is one-quarter of the MLD query interval. The number of queries, or the startup query count, is user configurable.

After startup, the MLD querier periodically sends MLD general queries at the MLD query interval to check for IPv6 multicast group members on the network. You can modify the query interval based on the actual condition of the network.

The MLDv1 querier sends MLD multicast-address-specific queries at the MLD last listener query interval when it receives an MLD done message. The MLDv2 querier sends MLD multicast-address-and-source-specific queries at the MLD last listener query interval when it receives a multicast group and multicast source mapping change report. The number of queries, or the last listener query count, equals the robustness variable (the maximum number of packet retransmissions).

A multicast listening host starts a timer for each IPv6 multicast group that it has joined when it receives an MLD query (general query, multicast-address-specific query, or multicast-address-and-source-specific query). The timer is initialized to a random value in the range of 0 to the maximum response delay advertised in the MLD query message. When the timer decreases to 0, the host sends an MLD membership report message to the IPv6 multicast group.

Configuration guidelines

To speed up the response of hosts to MLD queries and avoid simultaneous timer expirations causing MLD report traffic bursts, you must properly set the maximum response delay.

- For MLD general queries, the maximum response delay is set by the **max-response-time** command.
- For MLD multicast-address-specific query and multicast-address-and-source-specific query messages, the maximum response delay equals the last listener query interval.

When multiple multicast routers exist on the same subnet, the MLD querier is responsible for sending MLD query messages. If a non-querier router receives no MLD query from the querier when the other querier present interval expires, it considers the querier as having failed and starts a new querier election. Otherwise, the non-querier resets the other querier present timer.

To avoid frequent MLD querier changes, set the other querier present interval greater than the MLD query interval.

To avoid incorrect multicast group member removals, set the MLD query interval greater than the maximum response delay for MLD general queries.

Configuration procedure

To configure MLD query and response parameters globally:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter MLD view.	mld	N/A
3. Configure the MLD querier's robustness variable.	robust-count <i>robust-value</i>	2 times by default. A higher robustness variable makes the MLD querier more robust but results in a longer IPv6 multicast group timeout time.
4. Configure the startup query interval.	startup-query-interval <i>interval</i>	By default, the startup query interval is 1/4 of the "MLD query interval".
5. Configure the startup query count.	startup-query-count <i>value</i>	By default, the startup query count is set to the MLD querier's robustness variable.

Step	Command	Remarks
6. Configure the MLD query interval.	timer query <i>interval</i>	125 seconds by default.
7. Configure the maximum response delay for MLD general query messages.	max-response-time <i>interval</i>	10 seconds by default.
8. Configure the MLD last listener query interval.	last-listener-query-interval <i>interval</i>	1 second by default.
9. Configure the MLD other querier present interval.	timer other-querier-present <i>interval</i>	By default, the other querier present interval is determined by the formula "Other querier present interval (in seconds) = [MLD query interval] × [MLD querier's robustness variable] + [maximum response delay for MLD general query] / 2".

To configure MLD query and response parameters on an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the MLD querier's robustness variable.	mld robust-count <i>robust-value</i>	2 times by default.
4. Configure the startup query interval.	mld startup-query-interval <i>interval</i>	By default, the startup query interval is 1/4 of the "MLD query interval".
5. Configure the startup query count.	mld startup-query-count <i>value</i>	By default, the startup query count is the same as the robustness variable.
6. Configure the MLD query interval.	mld timer query <i>interval</i>	125 seconds by default.
7. Configure the maximum response delay for MLD general query messages.	mld max-response-time <i>interval</i>	10 seconds by default.
8. Configure the MLD last listener query interval.	mld last-listener-query-interval <i>interval</i>	1 second by default.
9. Configure the MLD other querier present interval.	mld timer other-querier-present <i>interval</i>	By default, the other querier present interval is determined by the formula "Other querier present interval (in seconds) = [MLD query interval] × [MLD querier's robustness variable] + [maximum response delay for MLD general query] / 2".

Configuring MLD fast-leave processing

In some applications, such as ADSL dial-up networking, only one multicast receiver host is attached to a port of the MLD querier. To allow fast response to the MLD done messages of the host when it switches frequently from one IPv6 multicast group to another, you can enable MLD fast-leave processing on the MLD querier.

With fast-leave processing enabled, after receiving an MLD done message from a host, the MLD querier sends a leave notification to the upstream immediately without first sending MLD multicast-address-specific queries. In this way, the leave latency is reduced on one hand, and the network bandwidth is saved on the other hand.

Configuring MLD fast-leave processing globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter MLD view.	mld	N/A
3. Configure MLD fast-leave processing.	fast-leave [group-policy <i>acl6-number</i>]	Disabled by default.

Configuring MLD fast-leave processing on an interface

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure MLD fast-leave processing.	mld fast-leave [group-policy <i>acl6-number</i>]	Disabled by default.

NOTE:

The MLD fast-leave processing configuration is effective on Layer 3 interfaces other than VLAN interfaces, including Layer 3 Ethernet ports, Layer 3 aggregate interfaces, and Tunnel interfaces.

Enabling the MLD host tracking function

With the MLD host tracking function, the switch can record the information of the member hosts that are receiving IPv6 multicast traffic, including the host IPv6 address, running duration, and timeout time. You can monitor and manage the member hosts according to the recorded information.

Enabling the MLD host tracking function globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter MLD view.	mld	N/A
3. Configure the MLD host tracking function globally.	host-tracking	Disabled by default

Enabling the MLD host tracking function on an interface

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type interface-number</i>	N/A
3. Enable the MLD host tracking function on the interface.	mld host-tracking	Disabled by default

Setting the DSCP value for MLD messages

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter MLD view.	mld	N/A
3. Set the DSCP value for MLD messages.	dscp <i>dscp-value</i>	Optional. By default, the DSCP value in MLD messages is 48.

Configuring MLD SSM mapping

Because of some possible restrictions, some receiver hosts on an SSM network might run MLDv1. To provide SSM service support for these receiver hosts, you need to configure the MLD SSM mapping feature on the last hop router.

Configuration prerequisites

Before you configure the MLD SSM mapping feature, complete the following tasks:

- Enable IPv6 forwarding and configure an IPv6 unicast routing protocol so that all devices in the domain can be interoperable at the network layer.
- Configure basic MLD functions.

Enabling MLD SSM mapping

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type interface-number</i>	N/A
3. Enable the MLD SSM mapping feature.	mld ssm-mapping enable	Disabled by default

NOTE:

To ensure SSM service for all hosts on a subnet, regardless of the MLD version running on the hosts, enable MLDv2 on the interface that forwards IPv6 multicast traffic onto the subnet.

Configuring MLD SSM mappings

By performing this configuration multiple times, you can map an IPv6 multicast group to different IPv6 multicast sources.

If MLDv2 is enabled on a VLAN interface of a switch, and if a port in that VLAN is configured as a simulated host, the simulated host will send MLDv2 reports even if you did not specify an IPv6 multicast source when you configure simulated joining with the **mld-snooping host-join** command. In this case, the corresponding IPv6 multicast group will not be created based on the configured MLD SSM mappings. For more information about the **mld-snooping host-join** command, see *IP Multicast Command Reference*.

To configure an MLD SSM mapping:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter MLD view.	mld	N/A
3. Configure an MLD SSM mapping.	ssm-mapping <i>ipv6-group-address</i> <i>prefix-length</i> <i>ipv6-source-address</i>	No MLD mappings are configured by default.

Configuring MLD proxying

Configuration prerequisites

Before you configure the MLD proxying feature, complete the following tasks:

- Enable IPv6 forwarding and configure an IPv6 unicast routing protocol so that all devices in the domain are interoperable at the network layer.
- Enable IPv6 multicast routing.

Enabling MLD proxying

You can enable MLD proxying on the interface in the direction toward the root of the multicast forwarding tree to make the device serve as an MLD proxy.

Configuration guidelines

- Each device can have only one interface serving as the MLD proxy interface.
- You cannot enable MLD on interfaces with MLD proxying enabled. Moreover, only the **mld require-router-alert**, **mld send-router-alert**, and **mld version** commands can take effect on such interfaces.
- You cannot enable other IPv6 multicast routing protocols (such as IPv6 PIM-DM or IPv6 PIM-SM) on interfaces with MLD proxying enabled, or vice versa. However, the **source-lifetime**, **source-policy**, and **ssm-policy** commands configured in IPv6 PIM view can still take effect.
- You cannot enable MLD proxying on a VLAN interface with MLD snooping enabled, or vice versa.

Configuration procedure

To enable MLD proxying:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable the MLD proxying feature.	mld proxying enable	Disabled by default

Configuring IPv6 multicast forwarding on a downstream interface

Typically, to avoid duplicate multicast flows, only queriers can forward IPv6 multicast traffic. On MLD proxy devices, a downstream interface must be a querier in order to forward IPv6 multicast traffic to downstream hosts. If the interface has failed in the querier election, you must manually enable IPv6 multicast forwarding on this interface.

ⓘ IMPORTANT:

On a multi-access network with more than one MLD proxy devices, you cannot enable IPv6 multicast forwarding on any other non-querier downstream interface after one of the downstream interfaces of these MLD proxy devices has been elected as the querier. Otherwise, duplicate multicast flows might be received on the multi-access network.

To enable IPv6 multicast forwarding on a downstream interface

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable IPv6 multicast forwarding on a non-querier downstream interface.	mld proxying forwarding	Disabled by default

Displaying and maintaining MLD

Task	Command	Remarks
Display MLD group information.	display mld group [<i>ipv6-group-address</i> interface <i>interface-type interface-number</i>] [static verbose] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display Layer 2 port information about MLD groups.	display mld group port-info [vlan <i>vlan-id</i>] [slot <i>slot-number</i>] [verbose] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display information about the hosts tracked by MLD on an interface.	display mld host interface <i>interface-type interface-number</i> group <i>ipv6-group-address</i> [source <i>ipv6-source-address</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about the hosts tracked by MLD on the Layer 2 ports.	display mld host port-info vlan <i>vlan-id</i> group <i>ipv6-group-address</i> [source <i>ipv6-source-address</i>] [vlan <i>vlan-id</i>] [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display MLD configuration and running information on the specified interface or all MLD-enabled interfaces.	display mld interface [<i>interface-type interface-number</i>] [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the information of the MLD proxying groups.	display mld proxying group [<i>group-address</i>] [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the information of the MLD routing table.	display mld routing-table [<i>ipv6-source-address</i> [<i>prefix-length</i>] <i>ipv6-group-address</i> [<i>prefix-length</i>] flags { act suc }] * [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display MLD SSM mappings.	display mld ssm-mapping <i>ipv6-group-address</i> [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the IPv6 multicast group information created based on the configured MLD SSM mappings.	display mld ssm-mapping group [<i>ipv6-group-address</i> interface <i>interface-type interface-number</i>] [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about the hosts that join based on the MLD SSM mappings on an interface.	display mld ssm-mapping host interface <i>interface-type</i> <i>interface-number</i> group <i>ipv6-group-address</i> source <i>ipv6-source-address</i> [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Remove the dynamic group entries of a specified MLD group or all MLD groups.	reset mld group { all interface <i>interface-type interface-number</i> { all <i>ipv6-group-address</i> [<i>prefix-length</i>] [<i>ipv6-source-address</i> [<i>prefix-length</i>]] } }	Available in user view

Task	Command	Remarks
Remove the dynamic Layer 2 port entries of a specified MLD group or all MLD groups.	reset mld group port-info { all <i>ipv6-group-address</i> } [vlan <i>vlan-id</i>]	Available in user view
Clear MLD SSM mappings.	reset mld ssm-mapping group { all interface <i>interface-type</i> <i>interface-number</i> { all <i>ipv6-group-address</i> [<i>prefix-length</i>] [<i>ipv6-source-address</i> [<i>prefix-length</i>]] } }	Available in user view

! **IMPORTANT:**

The **reset mld group** command might cause an interruption of receivers' reception of multicast data.

The **reset mld group** command cannot remove dynamic MLD group entries.

The **reset mld group port-info** command cannot remove the Layer 2 port entries of MLD groups.

The **display mld host interface** command can display information about the hosts tracked by MLD on Layer 3 interfaces other than VLAN interfaces.

The **display mld ssm-mapping host interface** command can display information about the hosts that join the group based on MLD SSM mappings on Layer 3 interfaces other than VLAN interfaces.

MLD configuration examples

Basic MLD functions configuration example

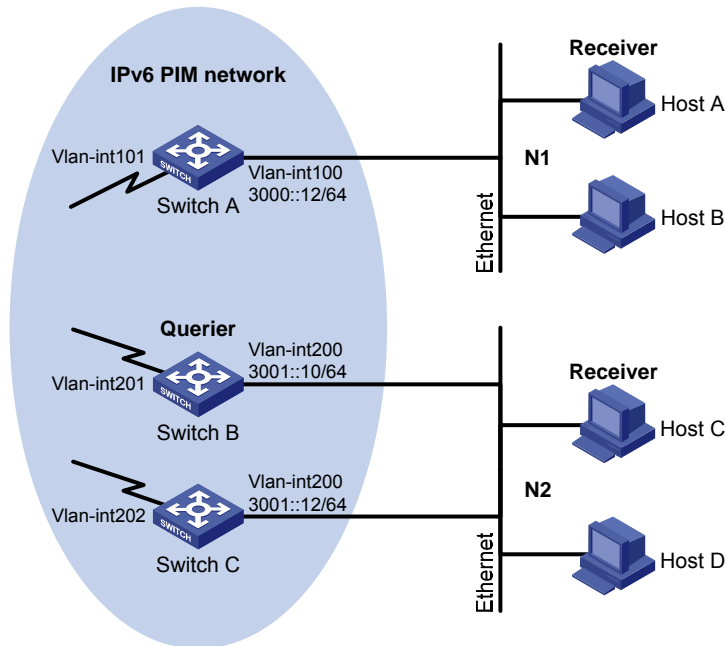
Network requirements

As shown in [Figure 87](#), receivers receive VOD information in the multicast mode. Receivers of different organizations form stub networks N1 and N2, and Host A and Host C are multicast receivers in N1 and N2 respectively.

MLDv1 runs between Switch A and N1. MLDv1 runs between the other two switches (Switch B and Switch C) and N2. Switch B acts as the MLD querier because it has a lower IPv6 address.

The hosts in N1 can join only IPv6 multicast group FF1E::101, and the hosts in N2 can join any IPv6 multicast groups.

Figure 87 Network diagram



Configuration procedure

1. Enable IPv6 forwarding and configure IPv6 addresses and IPv6 unicast routing:

Enable IPv6 forwarding on each switch and configure an IP address and prefix length for each interface as shown in [Figure 87](#). (Details not shown.)

Configure OSPFv3 for interoperation between the switches. Ensure the network-layer interoperation among the switches on the IPv6 PIM network and dynamic update of routing information between the switches through a unicast routing protocol. (Details not shown.)
2. Enable the IPv6 multicast routing, and enable IPv6 PIM-DM and MLD:

Enable IPv6 multicast routing on Switch A, enable IPv6 PIM-DM on each interface, and enable MLD on VLAN-interface 100.

```
<SwitchA> system-view
[SwitchA] multicast ipv6 routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] mld enable
[SwitchA-Vlan-interface100] pim ipv6 dm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim ipv6 dm
[SwitchA-Vlan-interface101] quit
```

Enable IPv6 multicast routing on Switch B, enable IPv6 PIM-DM on each interface, and enable MLD on VLAN-interface 200.

```
<SwitchB> system-view
[SwitchB] multicast ipv6 routing-enable
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] mld enable
[SwitchB-Vlan-interface200] pim ipv6 dm
[SwitchB-Vlan-interface200] quit
```

```
[SwitchB] interface vlan-interface 201
[SwitchB-Vlan-interface201] pim ipv6 dm
[SwitchB-Vlan-interface201] quit
```

Enable IPv6 multicast routing on Switch C, enable IPv6 PIM-DM on each interface, and enable MLD on VLAN-interface 200.

```
<SwitchC> system-view
[SwitchC] multicast ipv6 routing-enable
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] mld enable
[SwitchC-Vlan-interface200] pim ipv6 dm
[SwitchC-Vlan-interface200] quit
[SwitchC] interface vlan-interface 202
[SwitchC-Vlan-interface202] pim ipv6 dm
[SwitchC-Vlan-interface202] quit
```

3. Configure an IPv6 multicast group filter:

Configure an IPv6 multicast group filter on Switch A, so that the hosts connected to VLAN-interface 100 can join IPv6 multicast group FF1E::101 only.

```
[SwitchA] acl ipv6 number 2001
[SwitchA-acl6-basic-2001] rule permit source ffe::101 128
[SwitchA-acl6-basic-2001] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] mld group-policy 2001
[SwitchA-Vlan-interface100] quit
```

4. Verify the configuration:

Use the **display mld interface** command to display the MLD configuration and running information on each switch interface. Example:

Display MLD information on VLAN-interface 200 of Switch B.

```
[SwitchB] display mld interface vlan-interface 200
Vlan-interface200(FE80::200:5EFF:FE66:5100):
  MLD is enabled
  Current MLD version is 1
  Value of query interval for MLD(in seconds): 125
  Value of other querier present interval for MLD(in seconds): 255
  Value of maximum query response time for MLD(in seconds): 10
  Querier for MLD: FE80::200:5EFF:FE66:5100 (this router)
Total 1 MLD Group reported
```

MLD SSM mapping configuration example

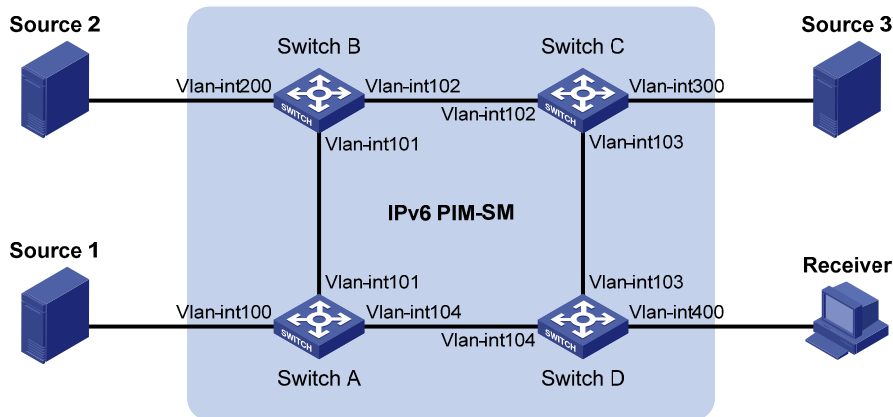
Network requirements

As shown in [Figure 88](#), the IPv6 PIM-SM domain applies both the ASM model and SSM model for IPv6 multicast delivery. Switch D's VLAN-interface 104 serves as the C-BSR and C-RP. The SSM group range is FF3E::/64.

MLDv2 runs on Switch D's VLAN-interface 400. The receiver host runs MLDv1, and does not support MLDv2. Therefore, the Receiver host cannot specify expected multicast sources in its membership reports.

Source 1, Source 2, and Source 3 send IPv6 multicast packets to multicast groups in the IPv6 SSM group range. You can configure the MLD SSM mapping feature on Switch D so that the receiver host will receive IPv6 multicast data from Source 1 and Source 3 only.

Figure 88 Network diagram



Device	Interface	IPv6 address	Device	Interface	IPv6 address
Source 1	—	1001::1/64	Source 3	—	3001::1/64
Source 2	—	2001::1/64	Receiver	—	4001::1/64
Switch A	Vlan-int100	1001::2/64	Switch C	Vlan-int300	3001::2/64
	Vlan-int101	1002::1/64		Vlan-int103	3002::1/64
	Vlan-int104	1003::1/64		Vlan-int102	2002::2/64
Switch B	Vlan-int200	2001::2/64	Switch D	Vlan-int400	4001::2/64
	Vlan-int101	1002::2/64		Vlan-int103	3002::2/64
	Vlan-int102	2002::1/64		Vlan-int104	1003::2/64

Configuration procedure

1. Enable IPv6 forwarding and configure IPv6 addresses and IPv6 unicast routing:
 Enable IPv6 forwarding on each switch and configure an IPv6 address and prefix length for each interface as shown in [Figure 88](#). (Details not shown.)
 Configure OSPFv3 for interoperability among the switches. Ensure the network-layer interoperation on the IPv6 PIM-SM domain and dynamic update of routing information among the switches through an IPv6 unicast routing protocol. (Details not shown.)
2. Enable IPv6 multicast routing, enable IPv6 PIM-SM on each interface and enable MLD and MLD SSM mapping on the host-side interface:

Enable IPv6 multicast routing on Switch D, enable IPv6 PIM-SM on each interface, and enable MLD (version 2) and MLD SSM mapping on VLAN-interface 400.

```
<SwitchD> system-view
[SwitchD] multicast ipv6 routing-enable
[SwitchD] interface vlan-interface 400
[SwitchD-Vlan-interface400] mld enable
[SwitchD-Vlan-interface400] mld version 2
[SwitchD-Vlan-interface400] mld ssm-mapping enable
[SwitchD-Vlan-interface400] pim ipv6 sm
[SwitchD-Vlan-interface400] quit
```

```
[SwitchD] interface vlan-interface 103
[SwitchD-Vlan-interface103] pim ipv6 sm
[SwitchD-Vlan-interface103] quit
[SwitchD] interface vlan-interface 104
[SwitchD-Vlan-interface104] pim ipv6 sm
[SwitchD-Vlan-interface104] quit
```

Enable IPv6 multicast routing on Switch A, and enable IPv6 PIM-SM on each interface.

```
<SwitchA> system-view
[SwitchA] multicast ipv6 routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] pim ipv6 sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim ipv6 sm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 104
[SwitchA-Vlan-interface104] pim ipv6 sm
[SwitchA-Vlan-interface104] quit
```

The configuration on Switch B and Switch C is similar to that on Switch A.

3. Configure a C-BSR and a C-RP:

Configure C-BSR and C-RP interfaces on Switch D.

```
[SwitchD] pim ipv6
[SwitchD-pim6] c-bsr 1003::2
[SwitchD-pim6] c-rp 1003::2
[SwitchD-pim6] quit
```

4. Configure the IPv6 SSM group range:

Configure the IPv6 SSM group range FF3E::/64 on Switch D.

```
[SwitchD] acl ipv6 number 2000
[SwitchD-acl6-basic-2000] rule permit source ff3e:: 64
[SwitchD-acl6-basic-2000] quit
[SwitchD] pim ipv6
[SwitchD-pim6] ssm-policy 2000
[SwitchD-pim6] quit
```

The configuration on Switch A, Switch B and Switch C is similar to that on Switch D.

5. Configure MLD SSM mappings:

Configure MLD SSM mappings on Switch D.

```
[SwitchD] mld
[SwitchD-mld] ssm-mapping ff3e:: 64 1001::1
[SwitchD-mld] ssm-mapping ff3e:: 64 3001::1
[SwitchD-mld] quit
```

6. Verify the configuration:

Display MLD SSM mapping information about the IPv6 multicast group FF3E::101 on Switch D.

```
[SwitchD] display mld ssm-mapping ff3e::101
Group: FF3E::101
Source list:
    1001::1
```

```

3001::1
# Display the IPv6 multicast group information created based on the configured MLD SSM
mappings on Switch D.
[SwitchD] display mld ssm-mapping group
Total 1 MLD SSM-mapping Group(s).
Interface group report information
Vlan-interface400 (4001::2):
  Total 1 MLD SSM-mapping Group reported
  Group Address: FF3E::101
  Last Reporter: 4001::1
  Uptime: 00:02:04
  Expires: off
# Display the IPv6 PIM routing table information on Switch D.
[SwitchD] display pim ipv6 routing-table
Total 0 (*, G) entry; 2 (S, G) entry

(1001::1, FF3E::101)
  Protocol: pim-ssm, Flag:
  UpTime: 00:13:25
  Upstream interface: Vlan-interface104
    Upstream neighbor: 1003::1
    RPF prime neighbor: 1003::1
  Downstream interface(s) information:
    Total number of downstreams: 1
    1: Vlan-interface400
      Protocol: mld, UpTime: 00:13:25, Expires: -

(3001::1, FF3E::101)
  Protocol: pim-ssm, Flag:
  UpTime: 00:13:25
  Upstream interface: Vlan-interface103
    Upstream neighbor: 3002::1
    RPF prime neighbor: 3002::1
  Downstream interface(s) information:
    Total number of downstreams: 1
    1: Vlan-interface400
      Protocol: mld, UpTime: 00:13:25, Expires: -

```

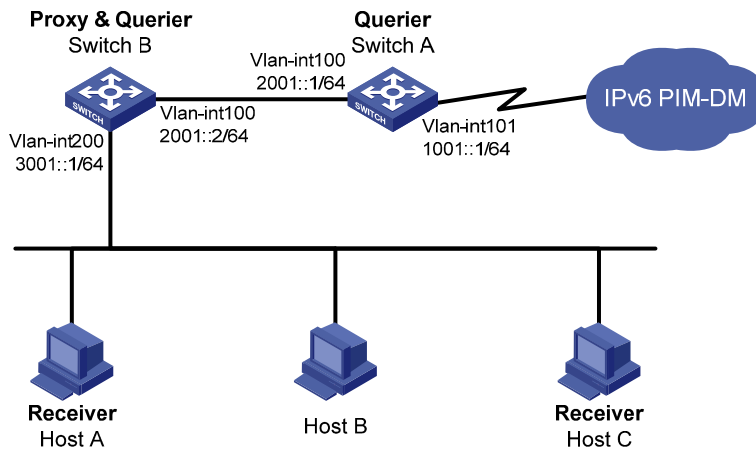
MLD proxying configuration example

Network requirements

As shown in [Figure 89](#), IPv6 PIM-DM runs on the core network. Host A and Host C in the stub network receive VOD information destined to multicast group FF3E::101.

Configure the MLD proxying feature on Switch B so that Switch B can maintain group memberships and forward IPv6 multicast traffic without running IPv6 PIM-DM.

Figure 89 Network diagram



Configuration procedure

1. Enable IPv6 forwarding and configure the IPv6 addresses:
Enable IPv6 forwarding on each switch and configure the IPv6 address and prefix length of each interface as per [Figure 89](#). (Details not shown.)
2. Enable IPv6 multicast routing, IPv6 PIM-DM, MLD, and MLD proxying:
Enable IPv6 multicast routing on Switch A, IPv6 PIM-DM on VLAN-interface 101, and MLD on VLAN-interface 100.

```
<SwitchA> system-view
[SwitchA] multicast ipv6 routing-enable
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim ipv6 dm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] mld enable
[SwitchA-Vlan-interface100] pim ipv6 dm
[SwitchA-Vlan-interface100] quit
```

Enable IPv6 multicast routing on Switch B, MLD proxying on VLAN-interface 100, and MLD on VLAN-interface 200.

```
<SwitchB> system-view
[SwitchB] multicast ipv6 routing-enable
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] mld proxying enable
[SwitchB-Vlan-interface100] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] mld enable
[SwitchB-Vlan-interface200] quit
```

3. Verify the configuration:
Display MLD configuration and operation information on VLAN-interface 100 of Switch B.

```
[SwitchB] display mld interface vlan-interface 100 verbose
Vlan-interface100(2001::2):
  MLD proxy is enabled
  Current MLD version is 1
  Multicast routing on this interface: enabled
```

```

Require-router-alert: disabled
# Display MLD group information on Switch A.
[SwitchA] display mld group
Total 1 MLD Group(s).
Interface group report information
Vlan-interface100(2001::1):
  Total 1 MLD Groups reported
  Group Address      Last Reporter      Uptime           Expires
  ff3e::101          2001::2            00:02:04         00:01:15

```

The output shows that the MLD reports sent from the hosts are forwarded to Switch A through the proxy interface, VLAN-interface 100 of Switch B.

Troubleshooting MLD

No member information on the receiver-side router

Symptom

When a host sends a message to announce its joining IPv6 multicast group G, no member information of multicast group G exists on the immediate router.

Analysis

- The correctness of networking and interface connections and whether the protocol layer of the interface is up directly affect the generation of IPv6 group member information.
- IPv6 multicast routing must be enabled on the router and MLD must be enabled on the interface connecting to the host.
- If the MLD version on the router interface is lower than that on the host, the router will not be able to recognize the MLD report from the host.
- If the **mld group-policy** command has been configured on an interface, the interface cannot receive report messages that fail to pass filtering.

Solution

1. Check that the networking, interface connections, and IP address configuration are correct. Check the interface information with the **display mld interface** command. If no information is output, the interface is in an abnormal state. This is usually because you have configured the **shutdown** command on the interface, the interface is not properly connected, or the IPv6 address configuration is not correctly done.
2. Use the **display current-configuration** command to verify that the IPv6 multicast routing is enabled. If not, carry out the **multicast ipv6 routing-enable** command in system view to enable IPv6 multicast routing. In addition, enable MLD on the corresponding interface.
3. You can use the **display mld interface** command to verify that the MLD version on the interface is lower than that on the host.
4. Use the **display current-configuration interface** command to verify that no ACL rule has been configured to restrict the host from joining IPv6 multicast group G. If an IPv6 ACL is configured to restrict the host from joining IPv6 multicast group G, the ACL must be modified to allow IPv6 multicast group G to receive report messages.

Inconsistent memberships on routers on the same subnet

Symptom

Different memberships are maintained on different MLD routers on the same subnet.

Analysis

- A router running MLD maintains multiple parameters for each interface, and these parameters influence one another, forming very complicated relationships. Inconsistent MLD interface parameter configurations for routers on the same subnet will surely result in inconsistent MLD memberships.
- Two MLD versions are available. Although routers running different MLD versions are compatible with hosts, all routers on the same subnet must run the same MLD version. Inconsistent MLD versions running on routers on the same subnet will also lead to inconsistent MLD memberships.

Solution

1. Use the **display current-configuration** command to verify the MLD configuration information on the interface.
2. Use the **display mld interface** command on all routers on the same subnet to check the MLD timers for inconsistent configuration.
3. Use the **display mld interface** command to verify that the routers are running the same MLD version.

Configuring IPv6 PIM (available only on the HP 5500 EI)

Overview

Protocol Independent Multicast for IPv6 (IPv6 PIM) provides IPv6 multicast forwarding by leveraging IPv6 unicast static routes or IPv6 unicast routing tables generated by any IPv6 unicast routing protocol, such as RIPng, OSPFv3, IS-ISv6, or BGP4+. IPv6 PIM uses an IPv6 unicast routing table to perform reverse path forwarding (RPF) check to implement IPv6 multicast forwarding. Independent of the IPv6 unicast routing protocols running on the device, IPv6 multicast routing can be implemented as long as the corresponding IPv6 multicast routing entries are created through IPv6 unicast routes. IPv6 PIM uses the reverse path forwarding (RPF) mechanism to implement IPv6 multicast forwarding. When an IPv6 multicast packet arrives on an interface of the device, RPF check is performed on it. If the RPF check succeeds, the device creates the corresponding routing entry and forwards the packet. If the RPF check fails, the device discards the packet. For more information about RPF, see "[Configuring IPv6 multicast routing and forwarding \(available only on the HP 5500 EI\)](#)."

Based on the implementation mechanism, IPv6 PIM supports the following types:

- Protocol Independent Multicast–Dense Mode for IPv6 (IPv6 PIM-DM)
- Protocol Independent Multicast–Sparse Mode for IPv6 (IPv6 PIM-SM)
- Bidirectional Protocol Independent Multicast for IPv6 (IPv6 BIDIR-PIM)
- Protocol Independent Multicast Source-Specific Multicast for IPv6 (IPv6 PIM-SSM)

To facilitate description, a network comprising IPv6 PIM–supporting routers is referred to as an "IPv6 PIM domain" in this document.

The term "router" in this document refers to both routers and Layer 3 switches.

The term "interface" in the IPv6 PIM features refers to Layer 3 interfaces, including VLAN interfaces and route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).

IPv6 PIM-DM overview

IPv6 PIM-DM is a type of dense mode IPv6 multicast protocol. It uses the push mode for IPv6 multicast forwarding, and is suitable for small-sized networks with densely distributed IPv6 multicast members.

The basic implementation of IPv6 PIM-DM is as follows:

- IPv6 PIM-DM assumes that at least one IPv6 multicast group member exists on each subnet of a network. Therefore, IPv6 multicast data is flooded to all nodes on the network. Then, branches without IPv6 multicast forwarding are pruned from the forwarding tree, leaving only those branches that contain receivers. This flood-and-prune process takes place periodically. That is, pruned branches resume IPv6 multicast forwarding when the pruned state times out and then data is flooded again down these branches, and then the branches are pruned again.
- When a new receiver on a previously pruned branch joins an IPv6 multicast group, to reduce the join latency, IPv6 PIM-DM uses the graft mechanism to resume IPv6 multicast data forwarding to that branch.

In general, the IPv6 multicast forwarding path is a source tree. That is, it is a forwarding tree with the IPv6 multicast source as its "root" and IPv6 multicast group members as its "leaves." Because the source tree is the shortest path from the IPv6 multicast source to the receivers, it is also called "shortest path tree (SPT)."

The working mechanism of IPv6 PIM-DM is summarized as follows:

- Neighbor discovery
- SPT establishment
- Graft
- Assert

Neighbor discovery

In an IPv6 PIM domain, a PIM router discovers IPv6 PIM neighbors, maintains IPv6 PIM neighboring relationships with other routers, and builds and maintains SPTs by periodically multicasting IPv6 PIM hello messages to all other IPv6 PIM routers on the local subnet.

NOTE:

Every IPv6 PIM enabled interface on a router sends hello messages periodically and, therefore, learns the IPv6 PIM neighboring information pertinent to the interface.

SPT establishment

The process of constructing an SPT is the flood and prune process.

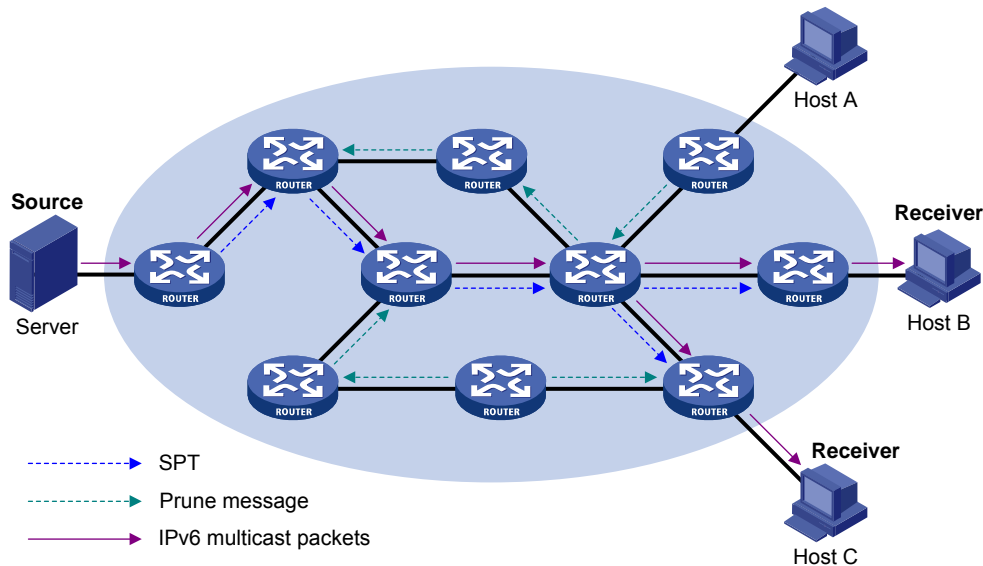
1. In an IPv6 PIM-DM domain, an IPv6 multicast source first floods IPv6 multicast packets when it sends IPv6 multicast data to IPv6 multicast group *G*. The packet undergoes an RPF check. If the packet passes the RPF check, the router creates an (*S*, *G*) entry and forwards the packet to all downstream nodes in the network. In the flooding process, an (*S*, *G*) entry is created on all the routers in the IPv6 PIM-DM domain.
2. The nodes without downstream receivers are pruned. A router that has no downstream receivers sends a prune message to the upstream node to notify the upstream node to delete the corresponding interface from the outgoing interface list in the (*S*, *G*) entry and stop forwarding subsequent packets addressed to that IPv6 multicast group down to this node.

An (*S*, *G*) entry contains the multicast source address *S*, IPv6 multicast group address *G*, outgoing interface list, and incoming interface.

For a given IPv6 multicast stream, the interface that receives the IPv6 multicast stream is referred to as "upstream," and the interfaces that forward the IPv6 multicast stream are referred to as "downstream."

A leaf router first initiates a prune process. As shown in [Figure 90](#), a router without any receiver attached to it (the router connected with Host A, for example) sends a prune message, and this prune process continues until only necessary branches remain in the IPv6 PIM-DM domain. These branches constitute the SPT.

Figure 90 SPT establishment in an IPv6 PIM-DM domain



The flood-and-prune process takes place periodically. A pruned state timeout mechanism is provided. A pruned branch restarts multicast forwarding when the pruned state times out and then is pruned again when it no longer has any multicast receiver.

NOTE:

Pruning has a similar implementation in IPv6 PIM-SM.

Graft

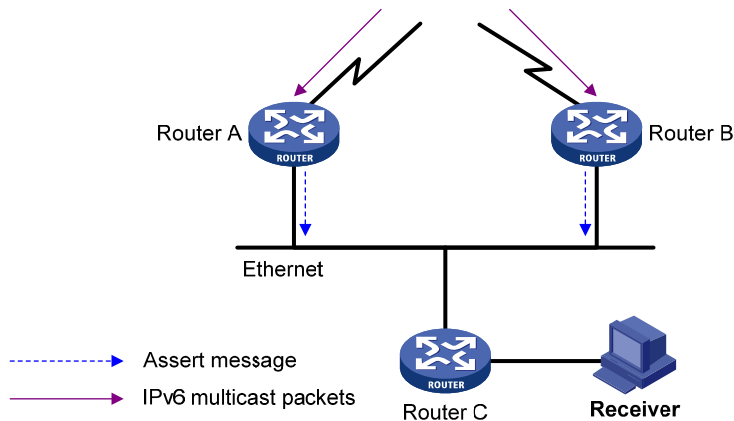
When a host attached to a pruned node joins an IPv6 multicast group, to reduce the join latency, IPv6 PIM-DM uses the graft mechanism to resume IPv6 multicast data forwarding to that branch. The process is as follows:

1. The node that needs to receive IPv6 multicast data sends a graft message toward its upstream node as a request to join the SPT again.
2. After receiving this graft message, the upstream node puts the interface on which the graft was received into the forwarding state and responds with a graft-ack message to the graft sender.
3. If the node that sent a graft message does not receive a graft-ack message from its upstream node, it keeps sending graft messages at a configurable interval until it receives an acknowledgment from its upstream node.

Assert

Where more than one multicast routers exists, the assert mechanism shuts off duplicate IPv6 multicast flows onto the same multi-access network. It does this by electing a unique IPv6 multicast forwarder on the multi-access network.

Figure 91 Assert mechanism



As shown in Figure 91, after Router A and Router B receive an (S, G) IPv6 multicast packet from the upstream node, they both forward the packet to the local subnet. As a result, the downstream node Router C receives two identical multicast packets, and both Router A and Router B, on their own local interface, receive a duplicate IPv6 multicast packet that the other has forwarded. After detecting this condition, both routers send an assert message to all IPv6 PIM routers on the local subnet through the interface that received the packet. The assert message contains the multicast source address (S), the multicast group address (G), and the preference and metric of the IPv6 unicast route/IPv6 MBGP route/IPv6 multicast static route to the source. By comparing these parameters, either Router A or Router B becomes the unique forwarder of the subsequent (S, G) IPv6 multicast packets on the multi-access subnet. The comparison process is as follows:

1. The router with a higher preference to the source wins.
2. If both routers have the same preference to the source, the router with a smaller metric to the source wins.
3. If a tie exists in the route metric to the source, the router with a higher IPv6 link-local address wins.

IPv6 PIM-SM overview

IPv6 PIM-DM uses the flood-and-prune principle to build SPTs for IPv6 multicast data distribution. Although an SPT has the shortest path, it is built with a low efficiency. Therefore the PIM-DM mode is not suitable for large-sized and medium-sized networks.

IPv6 PIM-SM is a type of sparse-mode IPv6 multicast protocol. It uses the pull mode for IPv6 multicast forwarding, and is suitable for large-sized and medium-sized networks with sparsely and widely distributed IPv6 multicast group members.

The basic implementation of IPv6 PIM-SM is as follows:

- IPv6 PIM-SM assumes that no hosts need to receive IPv6 multicast data. In the IPv6 PIM-SM mode, routers must specifically request a particular IPv6 multicast stream before the data is forwarded to them. The core task for IPv6 PIM-SM to implement IPv6 multicast forwarding will build and maintain rendezvous point trees (RPTs). An RPT is rooted at a router in the IPv6 PIM domain as the common node, or rendezvous point (RP), through which the IPv6 multicast data travels along the RPT and reaches the receivers.
- When a receiver is interested in the IPv6 multicast data addressed to a specific IPv6 multicast group, the router connected to this receiver sends a join message to the RP corresponding to that IPv6 multicast group. The path along which the message goes hop by hop to the RP forms a branch of the RPT.

- When an IPv6 multicast source sends IPv6 multicast streams to an IPv6 multicast group, the source-side designated router (DR) first registers the multicast source with the RP by sending register messages to the RP by unicast until it receives a register-stop message from the RP. The arrival of a register message at the RP triggers the establishment of an SPT. The IPv6 multicast source sends subsequent IPv6 multicast packets along the SPT to the RP. After reaching the RP, the IPv6 multicast packet is duplicated and delivered to the receivers along the RPT.

NOTE:

IPv6 multicast traffic is duplicated only where the distribution tree branches, and this process automatically repeats until the IPv6 multicast traffic reaches the receivers.

The working mechanism of IPv6 PIM-SM is summarized as follows:

- Neighbor discovery
- DR election
- RP discovery
- Embedded RP
- RPT establishment
- IPv6 Multicast source registration
- Switchover to SPT
- Assert

Neighbor discovery

IPv6 PIM-SM uses the similar neighbor discovery mechanism as IPv6 PIM-DM does. For more information, see "[Neighbor discovery](#)."

DR election

IPv6 PIM-SM also uses hello messages to elect a DR for a multi-access network (such as a LAN). The elected DR will be the only IPv6 multicast forwarder on this multi-access network.

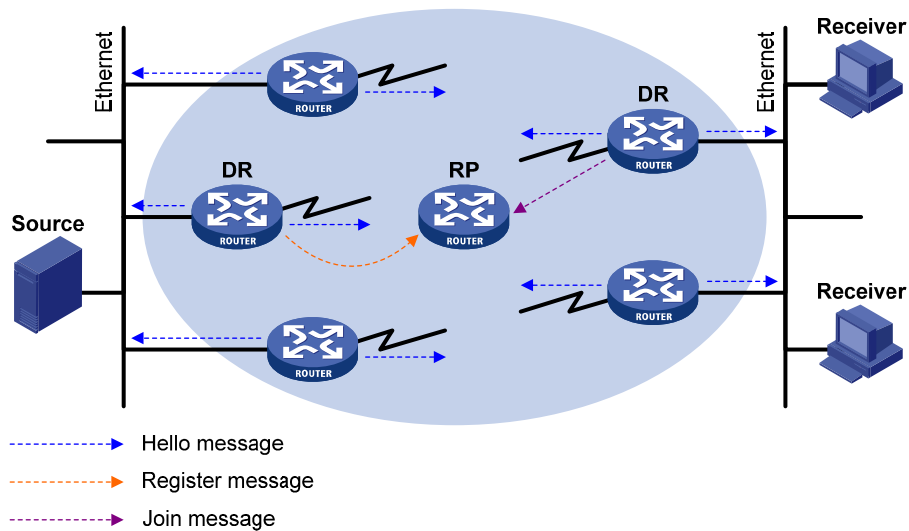
In the case of a multi-access network, a DR must be elected, no matter this network connects to IPv6 multicast sources or to receivers. The DR at the receiver side sends join messages to the RP; the DR at the IPv6 multicast source side sends register messages to the RP.

A DR is elected on a multi-access subnet by means of comparison of the priorities and IPv6 link-local addresses carried in hello messages.

MLD must be enabled on a device that acts as a receiver-side DR before receivers attached to this device can join IPv6 multicast groups through this DR.

For more information about MLD, see "[Configuring MLD \(available only on the HP 5500 EI\)](#)."

Figure 92 DR election



As shown in Figure 92, the DR election process is as follows:

1. Routers on the multi-access network send hello messages to one another. The hello messages contain the router priority for DR election. The router with the highest DR priority will become the DR.
2. In the case of a tie in the router priority, or if any router in the network does not support carrying the DR-election priority in hello messages, the router with the highest IPv6 link-local address will win the DR election.

When the DR works abnormally, a timeout in receiving hello message triggers a new DR election process among the other routers.

RP discovery

The RP is the core of an IPv6 PIM-SM domain. For a small-sized, simple network, one RP is enough for forwarding IPv6 multicast information throughout the network, and the position of the RP can be statically specified on each router in the IPv6 PIM-SM domain. In most cases, however, an IPv6 PIM-SM network covers a wide area and a huge amount of IPv6 multicast traffic must be forwarded through the RP. To lessen the RP burden and optimize the topological structure of the RPT, you can configure multiple candidate-RPs (C-RPs) in an IPv6 PIM-SM domain. Among them, an RP is dynamically elected through the bootstrap mechanism. Each elected RP serves a different multicast group range. For this purpose, you must configure a bootstrap router (BSR). The BSR serves as the administrative core of the IPv6 PIM-SM domain. An IPv6 PIM-SM domain can have only one BSR, but can have multiple candidate-BSRs (C-BSRs). If the BSR fails, a new BSR is automatically elected from the C-BSRs to avoid service interruption.

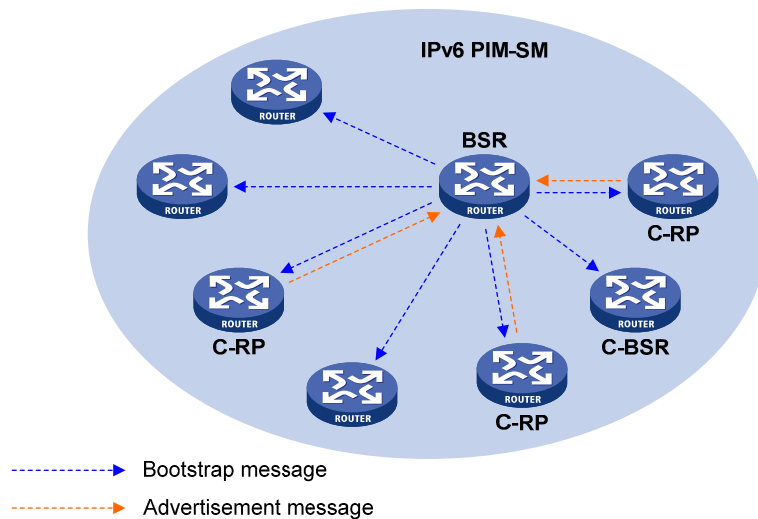
NOTE:

- An RP can serve IPv6 multiple multicast groups or all IPv6 multicast groups. Only one RP can serve a given IPv6 multicast group at a time.
 - A device can serve as a C-RP and a C-BSR at the same time.
-

As shown in Figure 93, each C-RP periodically unicasts its advertisement messages (C-RP-Adv messages) to the BSR. A C-RP-Adv message contains the address of the advertising C-RP and the IPv6 multicast group range it serves. The BSR collects these advertisement messages and chooses the appropriate C-RP information for each multicast group to form an RP-set, which is a database of mappings between IPv6

multicast groups and RPs. The BSR then encapsulates the RP-set in the bootstrap messages it periodically originates and floods the bootstrap messages (BSMs) to the entire IPv6 PIM-SM domain.

Figure 93 BSR and C-RPs



Based on the information in the RP-sets, all routers in the network can calculate the location of the corresponding RPs based on the following rules:

1. The C-RP with the highest priority wins.
2. If all the C-RPs have the same priority, their hash values are calculated through the hashing algorithm. The C-RP with the largest hash value wins.
3. If all the C-RPs have the same priority and hash value, the C-RP that has the highest IP address wins.

The hashing algorithm used for RP calculation is "Value (G, M, C_i) = (1103515245 * ((1103515245 * (G & M) + 12345) XOR C_i) + 12345) mod 2³¹."

Table 11 Values in the hashing algorithm

Value	Description
Value	Hash value.
G	The digest from the exclusive-or (XOR) operation between the 32-bit segments of the IPv6 multicast group address. For example, if the IPv6 multicast address is FFOE:C20:1A3:63::101, G = 0xFF0E0C20 XOR 0x01A30063 XOR 0x00000000 XOR 0x00000101.
M	Hash mask length.
C _i	The digest from the exclusive-or (XOR) operation between the 32-bit segments of the C-RP IPv6 address. For example, if the IPv6 address of the C-RP is 3FFE:B00:C18:1::10, C _i = 0x3FFE0B00 XOR 0x0C180001 XOR 0x00000000 XOR 0x00000010.
&	Logical operator of "and."
XOR	Logical operator of "exclusive-or."
mod	Modulo operator, which gives the remainder of an integer division.

Embedded RP

The embedded RP mechanism enables a router to resolve the RP address from an IPv6 multicast address so that the IPv6 multicast group is mapped to an RP. This RP can take the place of the statically configured RP or the RP dynamically calculated based on the BSR mechanism. The DR does not need to identify the RP address beforehand. The specific process is as follows.

At the receiver side, the following occur:

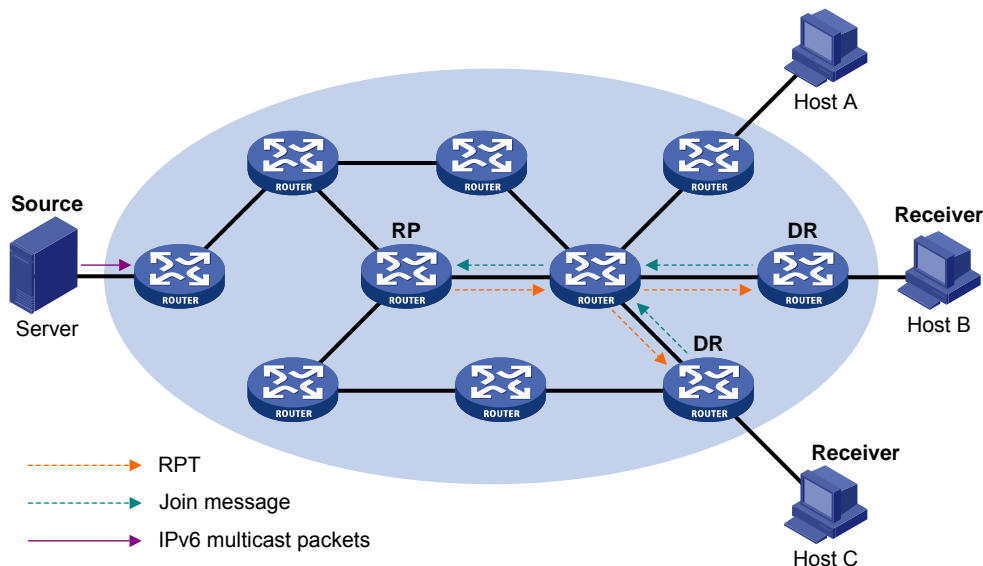
1. A receiver host initiates an MLD report to announce that it is joining an IPv6 multicast group.
2. After receiving the MLD report, the receiver-side DR resolves the RP address embedded in the IPv6 multicast address and sends a join message to the RP.

At the IPv6 multicast source side, the following occur:

1. The IPv6 multicast source sends IPv6 multicast traffic to the IPv6 multicast group.
2. The source-side DR resolves the RP address embedded in the IPv6 multicast address, and sends a register message to the RP.

RPT establishment

Figure 94 RPT establishment in an IPv6 PIM-SM domain



As shown in Figure 94, the process of building an RPT is as follows:

1. When a receiver joins IPv6 multicast group G, it uses an MLD report message to inform the directly connected DR.
2. After getting the IPv6 multicast group G's receiver information, the DR sends a join message, which is forwarded hop by hop to the RP that corresponds to the multicast group.
3. The routers along the path from the DR to the RP form an RPT branch. Each router on this branch generates a (*, G) entry in its forwarding table. The asterisk means any IPv6 multicast source. The RP is the root of the RPT, and the DRs are the leaves of the RPT.

The IPv6 multicast data addressed to the IPv6 multicast group G flows through the RP, reaches the corresponding DR along the established RPT, and finally is delivered to the receiver.

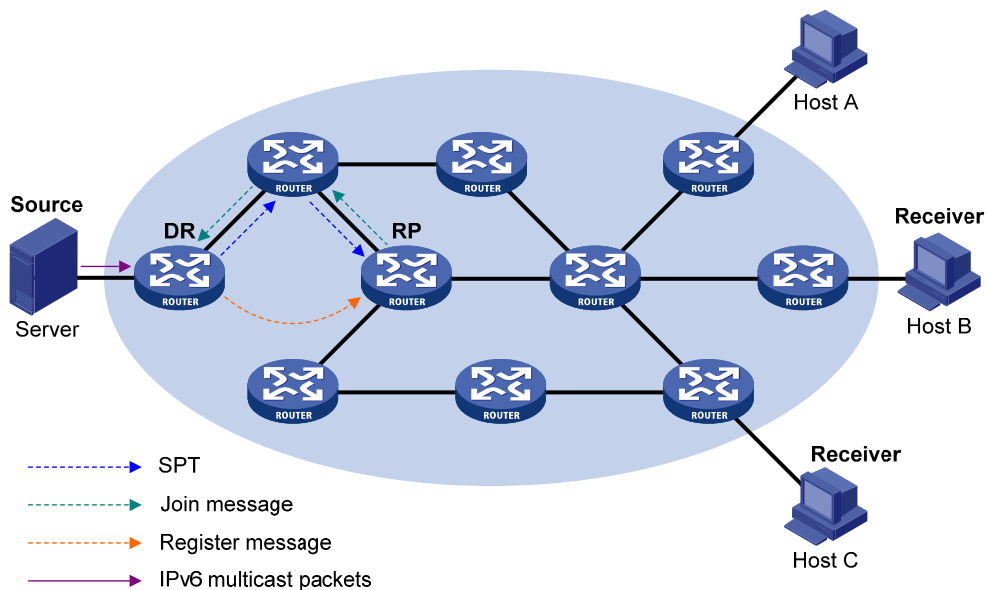
When a receiver is no longer interested in the IPv6 multicast data addressed to a multicast group G, the directly connected DR sends a prune message, which goes hop by hop along the RPT to the RP. After receiving the prune message, the upstream node deletes the interface connected with this downstream

node from the outgoing interface list and determines whether it has receivers for that IPv6 multicast group. If not, the router continues to forward the prune message to its upstream router.

Multicast source registration

The purpose of IPv6 multicast source registration will inform the RP about the existence of the IPv6 multicast source.

Figure 95 IPv6 multicast source registration



As shown in [Figure 95](#), the IPv6 multicast source registers with the RP as follows:

1. The IPv6 multicast source S sends the first IPv6 multicast packet to IPv6 multicast group G . After receiving the multicast packet, the DR that directly connects to the multicast source encapsulates the packet in a register message. Then it sends the message to the corresponding RP by unicast.
2. When the RP receives the register message, it extracts the multicast packet from the register message and forwards the multicast IPv6 multicast packet down the RPT, and sends an (S, G) join message hop by hop toward the IPv6 multicast source. The routers along the path from the RP to the IPv6 multicast source form an SPT branch. Each router on this branch generates an (S, G) entry in its forwarding table. The source-side DR is the root of the SPT, and the RP is the leaf of the SPT.
3. The subsequent IPv6 multicast data from the IPv6 multicast source travels along the established SPT to the RP, and then the RP forwards the data along the RPT to the receivers. When the IPv6 multicast traffic arrives at the RP along the SPT, the RP sends a register-stop message to the source-side DR by unicast to stop the source registration process.

NOTE:

The RP is configured to initiate an SPT switchover as described in this section. Otherwise, the DR at the IPv6 multicast source side keeps encapsulating multicast data in register messages and the registration process will not stop unless no outgoing interfaces exist in the (S, G) entry on the RP.

Switchover to SPT

In an IPv6 PIM-SM domain, an IPv6 multicast group corresponds to one RP and one RPT. Before the SPT switchover occurs, the DR at the IPv6 multicast source side encapsulates all multicast data destined to the multicast group in register messages and sends these messages to the RP. After receiving these register messages, the RP extracts the multicast data and sends the multicast data down the RPT to the DRs at the

receiver side. The RP acts as a transfer station for all IPv6 multicast packets. The whole process involves the following issues:

- The DR at the source side and the RP need to implement complicated encapsulation and de-encapsulation of IPv6 multicast packets.
- IPv6 multicast packets are delivered along a path that might not be the shortest one.
- An increase in IPv6 multicast traffic heavily burdens the RP, increasing the risk of failure.

To solve the issues, IPv6 PIM-SM allows an RP or the DR at the receiver side to initiate an SPT switchover process:

1. The RP initiates an SPT switchover process.

The RP can periodically check the passing-by IPv6 multicast packets. If it finds that the traffic rate exceeds a configurable threshold, the RP sends an (S, G) join message hop by hop toward the IPv6 multicast source to establish an SPT between the DR at the source side and the RP. Subsequent IPv6 multicast data travels along the established SPT to the RP.

For more information about the SPT switchover initiated by the RP, see "[Multicast source registration](#)."

2. The receiver-side DR initiates an SPT switchover process

After receiving the first IPv6 multicast packet, the receiver-side DR initiates an SPT switchover process, as follows:

- The receiver-side DR sends an (S, G) join message hop by hop toward the IPv6 multicast source. When the join message reaches the source-side DR, all the routers on the path have installed the (S, G) entry in their forwarding table, and thus an SPT branch is established.
- When the IPv6 multicast packets travel to the router where the RPT and the SPT deviate, the router drops the multicast packets received from the RPT and sends an RP-bit prune message hop by hop to the RP. After receiving this prune message, the RP sends a prune message toward the IPv6 multicast source (suppose only one receiver exists) to implement SPT switchover.
- IPv6 multicast data is directly sent from the source to the receivers along the SPT.

IPv6 PIM-SM builds SPTs through SPT switchover more economically than IPv6 PIM-DM does through the flood-and-prune mechanism.

Assert

IPv6 PIM-SM uses a similar assert mechanism as IPv6 PIM-DM does. For more information, see "[Assert](#)."

IPv6 BIDIR-PIM overview

In some many-to-many applications, such as multi-side video conference, there might be multiple receivers interested in multiple IPv6 multicast sources simultaneously. With IPv6 PIM-DM or IPv6 PIM-SM, each router along the SPT must create an (S, G) entry for each IPv6 multicast source, consuming a lot of system resources. IPv6 BIDIR-PIM is introduced to address this problem. Derived from IPv6 PIM-SM, IPv6 BIDIR-PIM builds and maintains bidirectional RPTs, each of which is rooted at an RP and connects IPv6 multiple multicast sources with multiple receivers. Traffic from the IPv6 multicast sources is forwarded through the RP to the receivers along the bidirectional RPT. In this case, each router needs to maintain only a (*, G) multicast routing entry, saving system resources.

IPv6 BIDIR-PIM is suitable for networks with dense multicast sources and dense receivers.

The working mechanism of IPv6 BIDIR-PIM is summarized as follows:

- Neighbor discovery

- RP discovery
- DF election
- Bidirectional RPT building

Neighbor discovery

IPv6 BIDIR-PIM uses the same neighbor discovery mechanism as IPv6 PIM-SM does. For more information, see "[Neighbor discovery](#)."

RP discovery

IPv6 BIDIR-PIM uses the same RP discovery mechanism as IPv6 PIM-SM does. For more information, see "[RP discovery](#)."

In IPv6 PIM-SM, an RP must be specified with a real IPv6 address. In IPv6 BIDIR-PIM, however, an RP can be specified with a virtual IPv6 address, which is called the rendezvous point address (RPA). The link corresponding to the RPA's subnet is called the rendezvous point link (RPL). All interfaces connected to the RPL can act as RPs, which back up one another.

In IPv6 BIDIR-PIM, an RPF interface is the interface pointing to an RP, and an RPF neighbor is the address of the next hop to the RP.

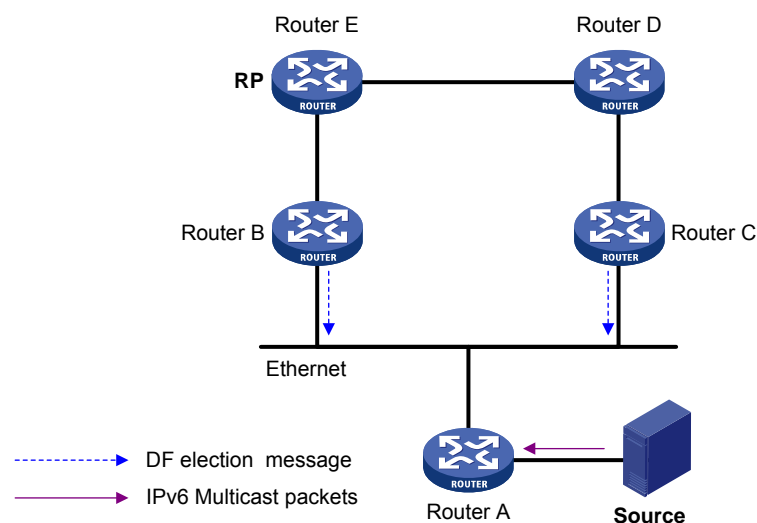
DF election

On a network segment with multiple multicast routers, the same multicast packets might be forwarded to the RP repeatedly. To address this issue, IPv6 BIDIR-PIM uses a DF election mechanism to elect a unique designated forwarder (DF) for each RP on every network segment within the IPv6 BIDIR-PIM domain, and allows only the DF to forward multicast data to the RP.

NOTE:

DF election is not necessary for an RPL.

Figure 96 DF election



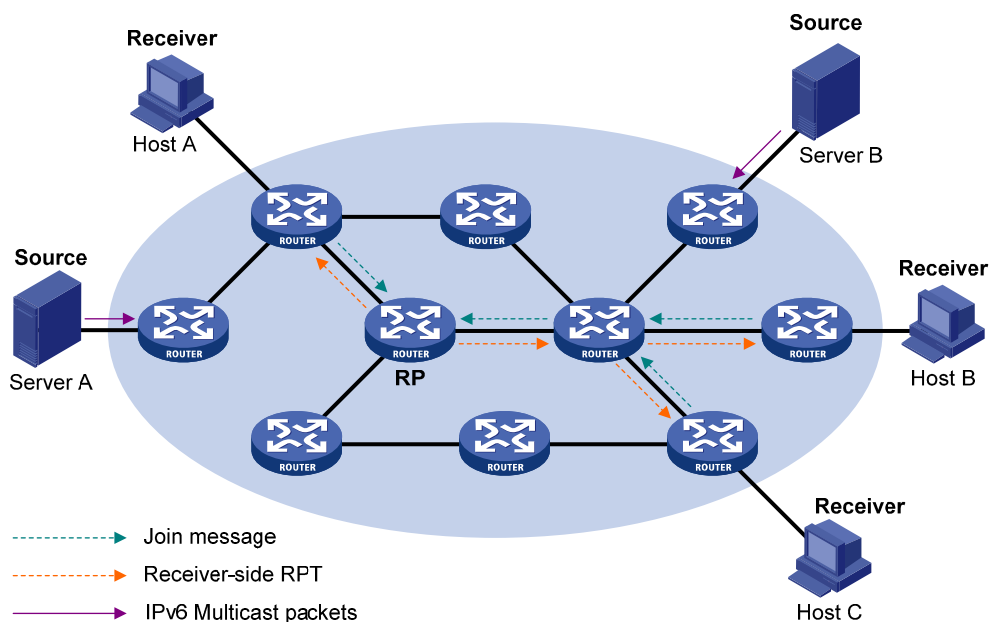
As shown in [Figure 96](#), without the DF election mechanism, both Router B and Router C can receive multicast packets from Router A, and they might both forward the packets to downstream routers on the local subnet. As a result, the RP (Router E) receives duplicate multicast packets. With the DF election mechanism, once receiving the RP information, Router B and Router C initiate a DF election process for the RP:

1. Router B and Router C multicast DF election messages to all PIM routers (224.0.0.13). The election messages carry the RP's address, and the priority and metric of the unicast route, MBGP route, or multicast static route to the RP.
2. The router with a route of the highest priority becomes the DF.
3. In the case of a tie, the router with the route with the lowest metric wins the DF election.
4. In the case of a tie in the metric, the router with the highest link-local IPv6 address wins.

Bidirectional RPT building

A bidirectional RPT comprises a receiver-side RPT and a source-side RPT. The receiver-side RPT is rooted at the RP and takes the routers directly connected with the receivers as leaves. The source-side RPT is also rooted at the RP but takes the routers directly connected with the IPv6 multicast sources as leaves. The processes for building these two parts are different.

Figure 97 RPT building at the receiver side

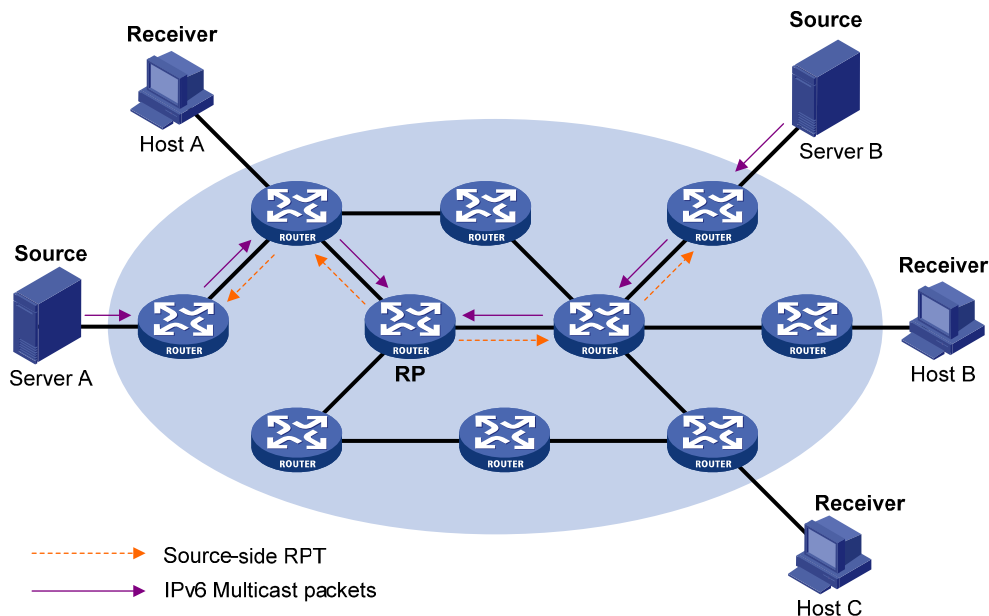


As shown in [Figure 97](#), the process for building a receiver-side RPT is similar to that for building an RPT in IPv6 PIM-SM:

1. When a receiver joins IPv6 multicast group G , it uses an MLD message to inform the directly connected router.
2. After getting the receiver information, the router sends a join message, which is forwarded hop by hop to the RP of the IPv6 multicast group.
3. The routers along the path from the receiver's directly connected router to the RP form an RPT branch, and each router on this branch adds a $(*, G)$ entry to its forwarding table. The $*$ means any IPv6 multicast source.

When a receiver is no longer interested in the multicast data addressed to IPv6 multicast group G , the directly connected router sends a prune message, which goes hop by hop along the reverse direction of the RPT to the RP. After receiving the prune message, each upstream node deletes the interface connected with the downstream node from the outgoing interface list and checks whether it has receivers in that IPv6 multicast group. If not, the router continues to forward the prune message to its upstream router.

Figure 98 RPT building at the multicast source side



As shown in Figure 98, the process of building a source-side RPT is relatively simple:

1. When an IPv6 multicast source sends IPv6 multicast packets to IPv6 multicast group G, the DF in each network segment unconditionally forwards the packets to the RP.
2. The routers along the path from the source's directly connected router to the RP form an RPT branch. Each router on this branch adds a (*, G) entry to its forwarding table. The * means any IPv6 multicast source.

After a bidirectional RPT is built, multicast traffic is forwarded along the source-side RPT and receiver-side RPT from IPv6 multicast sources to receivers.

NOTE:

If a receiver and an IPv6 multicast source are at the same side of the RP, the source-side RPT and the receiver-side RPT might meet at a node before reaching the RP. In this case, IPv6 multicast packets from the IPv6 multicast source to the receiver are directly forwarded by the node to the receiver, instead of by the RP.

IPv6 administrative scoping overview

Division of IPv6 PIM-SM domains

Typically, an IPv6 PIM-SM/IPv6 BIDIR-PIM domain contains only one BSR, which is responsible for advertising RP-set information within the entire IPv6 PIM-SM/IPv6 BIDIR-PIM domain. The information for all multicast groups is forwarded within the network scope administered by the BSR. We call this IPv6 non-scoped BSR mechanism.

To implement refined management, an IPv6 PIM-SM/IPv6 BIDIR-PIM domain can be divided into one IPv6 global scope zone and multiple IPv6 administratively scoped zones (IPv6 admin-scope zones). We call this IPv6 administrative scoping mechanism.

The IPv6 administrative scoping mechanism effectively releases stress on the management in a single-BSR domain and enables provision of zone-specific services using private group addresses.

IPv6 admin-scope zones correspond to IPv6 multicast groups with different scope values in their group addresses. The boundary of the IPv6 admin-scope zone is formed by zone border routers (ZBRs). Each IPv6 admin-scope zone maintains one BSR, which serves multicast groups within a specific scope. IPv6 multicast protocol packets, such as assert messages and bootstrap messages, for a specific group range cannot cross the IPv6 admin-scope zone boundary. IPv6 multicast group ranges served by different IPv6 admin-scope zones can overlap. An IPv6 multicast group is valid only within its local IPv6 admin-scope zone, functioning as a private group address.

The IPv6 global scope zone maintains a BSR, which serves the IPv6 multicast groups with the Scope field in their group addresses being 14.

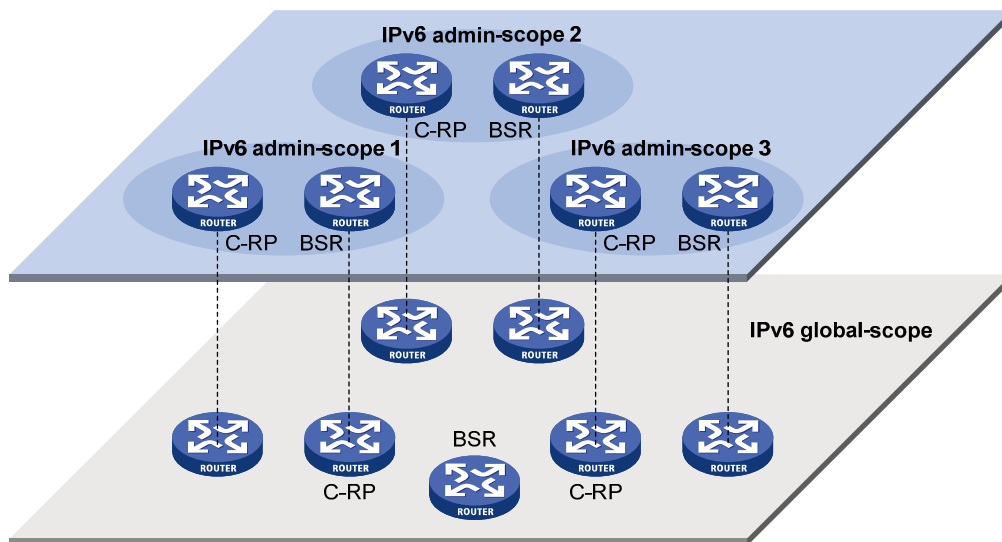
Relationship between IPv6 admin-scope zones and the IPv6 global scope zone

The IPv6 global scope zone and each IPv6 admin-scope zone have their own C-RPs and BSRs. These devices are effective only in their respective IPv6 admin-scope zones. That is, BSR election and RP election are implemented independently within each IPv6 admin-scope zone. Each IPv6 admin-scope zone has its own boundary. The multicast information cannot cross this border in either direction. A better understanding of the IPv6 global scope zone and IPv6 admin-scope zones should be based on geographical space and group address range.

- Geographical space

IPv6 admin-scope zones are logical zones specific to particular multicast groups. The multicast packets of these multicast groups are confined within the local IPv6 admin-scope zone and cannot cross the boundary of the zone.

Figure 99 Relationship between admin-scope zones and the global scope zone in geographic space

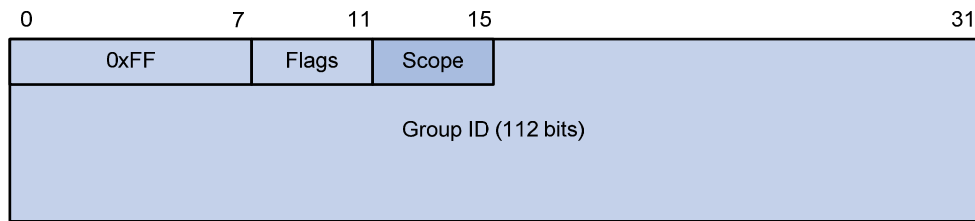


As shown in Figure 99, for multicast groups with the same Scope field in their group addresses, IPv6 admin-scope zones must be geographically separated from one another. Namely, a router must not serve different admin-scope zones. In other words, different admin-scope zones contain different routers, whereas the global scope zone covers all routers in the IPv6 PIM-SM/IPv6 BIDIR-PIM domain. Multicast packets that do not belong to any admin-scope zones can be transmitted in the entire IPv6 PIM-SM/IPv6 BIDIR-PIM domain.

- Multicast group address Scope field

As shown in Figure 100, the Scope field in each IPv6 multicast group address indicates the admin-scope zone the corresponding multicast group belongs to.

Figure 100 IPv6 multicast address format



The admin-scope zone range increases with the value of the Scope field. For example, value E indicates IPv6 global scope, which contains other admin-scope zones with the Scope field values smaller than E. Possible values of the Scope field are given in [Table 2](#).

Table 2 Values of the Scope field

Value	Meaning	Remarks
0, F	Reserved	N/A
1	Interface-local scope	N/A
2	Link-local scope	N/A
3	Subnet-local scope	IPv6 admin-scope zone
4	Admin-local scope	IPv6 admin-scope zone
5	Site-local scope	IPv6 admin-scope zone
6, 7, 9 through D	Unassigned	IPv6 admin-scope zone
8	Organization-local scope	IPv6 admin-scope zone
E	Global scope	IPv6 global-scope zone

IPv6 PIM-SSM overview

The source-specific multicast (SSM) model and the any-source multicast (ASM) model are opposites. The ASM model includes the IPv6 PIM-DM and IPv6 PIM-SM modes. You can implement the SSM model by leveraging part of the IPv6 PIM-SM technique. It is also called "IPv6 PIM-SSM."

The SSM model provides a solution for source-specific multicast. It maintains the relationships between hosts and routers through MLDv2.

In actual application, MLDv2 and part of IPv6 PIM-SM technique is adopted to implement the SSM model. In the SSM model, receivers know exactly where an IPv6 multicast source is located by using advertisements, consultancy, and so on. This model does not require RP or RPT, and it does not require a source registration process for the purpose of discovering IPv6 multicast sources in other IPv6 PIM domains.

In IPv6 PIM-SSM, the term "channel " refers to an IPv6 multicast group, and the term "channel subscription" refers to a join message.

The working mechanism of IPv6 PIM-SSM is summarized as follows:

- Neighbor discovery
- DR election
- SPT building

Neighbor discovery

IPv6 PIM-SSM uses the same neighbor discovery mechanism as in IPv6 PIM-SM. For more information, see "[Neighbor discovery](#)."

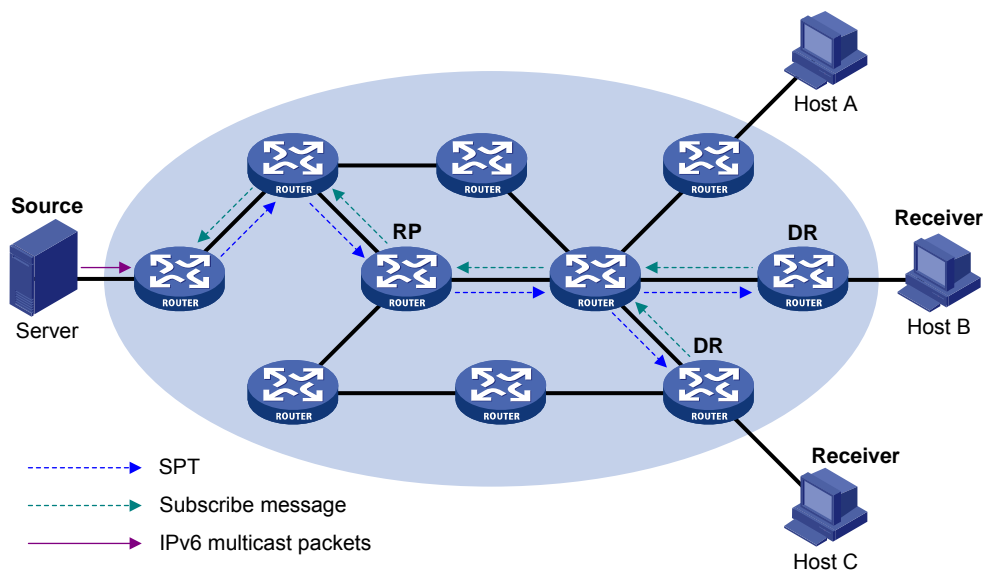
DR election

IPv6 PIM-SSM uses the same DR election mechanism as in IPv6 PIM-SM. For more information, see "[DR election](#)."

SPT building

The decision to build an RPT for IPv6 PIM-SM or an SPT for IPv6 PIM-SSM depends on whether the IPv6 multicast group that the receiver will join falls into the IPv6 SSM group range. The IPv6 SSM group range that IANA has reserved is FF3x::/32, where x represents any legal address scope.

Figure 101 Building an SPT in IPv6 PIM-SSM



As shown in [Figure 101](#), Hosts B and C are IPv6 multicast information receivers. They send an MLDv2 report message to the respective DRs to announce that they are interested in the information about the specific IPv6 multicast source S and that sent to the IPv6 multicast group G.

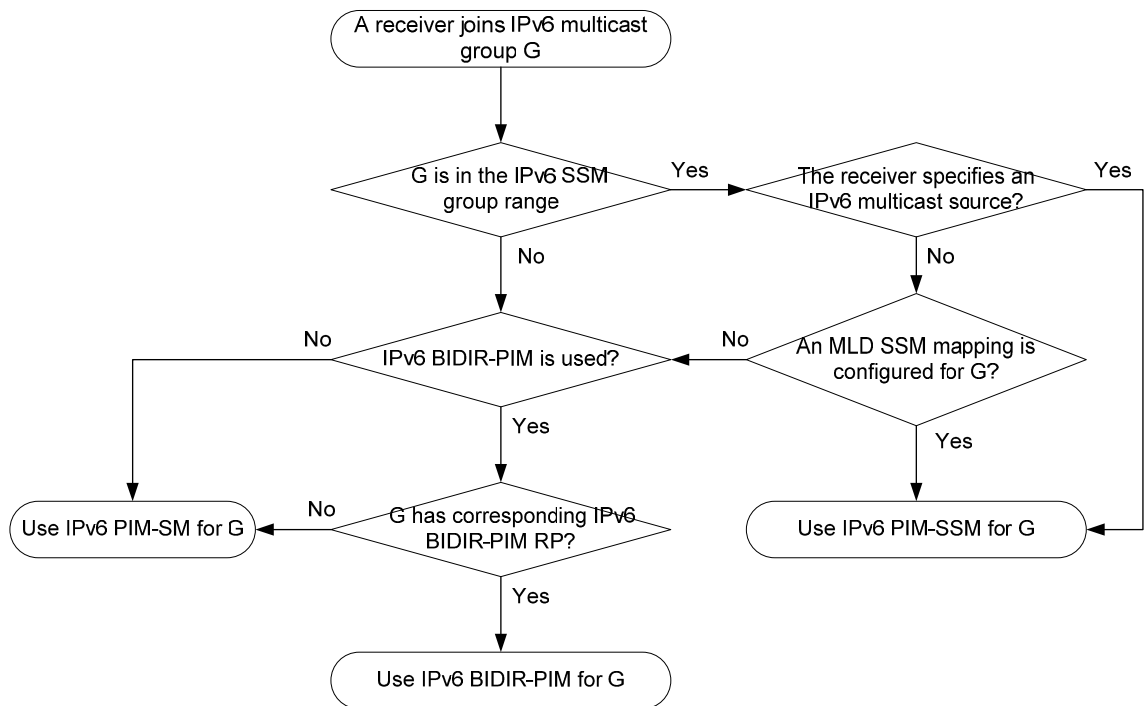
The DR that has received the report first determines whether the IPv6 group address in this message falls into the IPv6 SSM group range and then does the following:

- If the IPv6 group address in the message does fall into the IPv6 SSM group range, the IPv6 PIM-SSM model is built. The DR sends a channel subscription message hop by hop toward the IPv6 multicast source S. An (S, G) entry is created on all routers on the path from the DR to the source. Thus, an SPT is built in the network, with the source S as its root and receivers as its leaves. This SPT is the transmission channel in IPv6 PIM-SSM.
- If the IPv6 group address in the message does not fall into the IPv6 SSM group range, the DR follows the IPv6 PIM-SM process. The receiver-side DR sends a (*, G) join message to the RP, and the source-side DR registers the IPv6 multicast source.

Relationships among IPv6 PIM protocols

In an IPv6 PIM network, IPv6 PIM-DM cannot work with IPv6 PIM-SM, IPv6 BIDIR-PIM, or IPv6 PIM-SSM. However, IPv6 PIM-SM, IPv6 BIDIR-PIM, and IPv6 PIM-SSM can work together. When they work together, which one is chosen for a receiver trying to join a group depends, as shown in [Figure 102](#).

Figure 102 Relationships among IPv6 PIM protocols



For more information about MLD SSM mapping, see "[Configuring MLD \(available only on the HP 5500 EI\)](#)."

Protocols and standards

- RFC 3973, *Protocol Independent Multicast-Dense Mode(PIM-DM):Protocol Specification(Revised)*
- RFC 4601, *Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification (Revised)*
- RFC 3956, *Embedding the Rendezvous Point (RP) Address in an IPv6 Multicast Address*
- RFC 5015, *Bidirectional Protocol Independent Multicast (BIDIR-PIM)*
- RFC 5059, *Bootstrap Router (BSR) Mechanism for Protocol Independent Multicast (PIM)*
- RFC 4607, *Source-Specific Multicast for IP*
- draft-ietf-ssm-overview-05, *An Overview of Source-Specific Multicast (SSM)*

Configuring IPv6 PIM-DM

IPv6 PIM-DM configuration task list

Task	Remarks
Enabling IPv6 PIM-DM	Required
Enabling state-refresh capability	Optional
Configuring state refresh parameters	Optional
Configuring IPv6 PIM-DM graft retry period	Optional
Configuring IPv6 PIM common features	Optional

Configuration prerequisites

Before you configure IPv6 PIM-DM, complete the following tasks:

- Enable IPv6 forwarding and configure an IPv6 unicast routing protocol so that all devices in the domain are interoperable at the network layer.
- Determine the interval between state refresh messages.
- Determine the minimum time to wait before receiving a new refresh message.
- Determine the hop limit value of state-refresh messages.
- Determine the graft retry period.

Enabling IPv6 PIM-DM

With IPv6 PIM-DM enabled, a router sends hello messages periodically to discover IPv6 PIM neighbors and processes messages from the IPv6 PIM neighbors. When you deploy an IPv6 PIM-DM domain, enable IPv6 PIM-DM on all non-border interfaces of routers.

! IMPORTANT:

- All the interfaces of the same device must operate in the same IPv6 PIM mode.
- IPv6 PIM-DM cannot be used for IPv6 multicast groups in the IPv6 SSM group range.

To enable IPv6 PIM-DM:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IPv6 multicast routing.	multicast ipv6 routing-enable	Disabled by default.
3. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Enable IPv6 PIM-DM.	pim ipv6 dm	Disabled by default.

For more information about the **multicast ipv6 routing-enable** command, see *IP Multicast Command Reference*.

Enabling state-refresh capability

Pruned interfaces resume multicast forwarding when the pruned state times out. To prevent this, the router directly connected with the IPv6 multicast source periodically sends an (S, G) state-refresh message, which is forwarded hop by hop along the initial flooding path of the IPv6 PIM-DM domain, to refresh the

prune timer state of all the routers on the path. A multi-access subnet can have the state-refresh capability only if the state-refresh capability is enabled on all IPv6 PIM routers on the subnet.

To enable the state-refresh capability:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Enable the state-refresh capability.	pim ipv6 state-refresh-capable	Optional. Enabled by default.

Configuring state refresh parameters

The router directly connected with the multicast source periodically sends state-refresh messages. You can configure the interval for sending such messages.

A router might receive multiple state-refresh messages within a short time. Some messages might be duplicated messages. To keep a router from receiving such duplicated messages, you can configure the time that the router must wait before receiving the next state-refresh message. If the router receives a new state-refresh message within the waiting time, it discards it. If this timer times out, the router will accept a new state-refresh message, refresh its own IPv6 PIM-DM state, and reset the waiting timer.

The hop limit value of a state-refresh message decrements by 1 whenever it passes a router before it is forwarded to the downstream node until the hop limit value comes down to 0. In a small network, a state-refresh message might cycle in the network. To control the propagation scope of state-refresh messages, you must configure an appropriate hop limit value based on the network size.

Perform the following configurations on all routers in the IPv6 PIM domain.

To configure state-refresh parameters:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Configure the interval between state-refresh messages.	state-refresh-interval <i>interval</i>	Optional. 60 seconds by default.
4. Configure the time to wait before receiving a new state-refresh message.	state-refresh-rate-limit <i>interval</i>	Optional. 30 seconds by default.
5. Configure the hop limit value of state-refresh messages.	state-refresh-hoplimit <i>hoplimit-value</i>	Optional. 255 by default.

Configuring IPv6 PIM-DM graft retry period

In IPv6 PIM-DM, graft is the only type of message that uses the acknowledgment mechanism. In an IPv6 PIM-DM domain, if a router does not receive a graft-ack message from the upstream router within the specified time after it sends a graft message, the router keeps sending new graft messages at a

configurable interval (namely, graft retry period) until it receives a graft-ack message from the upstream router.

To configure the IPv6 PIM-DM graft retry period:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the graft retry period.	pim ipv6 timer graft-retry <i>interval</i>	Optional. 3 seconds by default.

For more information about the configuration of other timers in IPv6 PIM-DM, see "[Configuring IPv6 PIM common timers](#)."

Configuring IPv6 PIM-SM

IPv6 PIM-SM configuration task list

Task	Remarks	
Enabling IPv6 PIM-SM	Required.	
Configuring an RP	Configuring a static RP	Required.
	Configuring a C-RP	Use any approach.
	Enabling embedded RP	
Configuring a BSR	Configuring C-RP timers globally	Optional.
	Configuring a C-BSR	Required.
	Configuring an IPv6 PIM domain border	Optional.
	Configuring C-BSR parameters globally	Optional.
	Configuring C-BSR timers	Optional.
Configuring IPv6 administrative scoping	Disabling BSM semantic fragmentation	Optional.
	Enabling IPv6 administrative scoping	Optional.
	Configuring an IPv6 admin-scope zone boundary	Optional.
Configuring IPv6 multicast source registration	Optional.	
Configuring C-BSRs for IPv6 admin-scope zones	Optional.	
Disabling SPT switchover	Optional.	
Configuring IPv6 PIM common features	Optional.	

Configuration prerequisites

Before you configure IPv6 PIM-SM, complete the following tasks:

- Enable IPv6 forwarding and configure an IPv6 unicast routing protocol so that all devices in the domain are interoperable at the network layer.
- Determine the IP address of a static RP and the ACL rule defining the range of IPv6 multicast groups to be served by the static RP.
- Determine the C-RP priority and the ACL rule defining the range of IPv6 multicast groups to be served by each C-RP.
- Determine the legal C-RP address range and the ACL rule defining the range of IPv6 multicast groups to be served.
- Determine the C-RP-Adv interval.
- Determine the C-RP timeout.
- Determine the C-BSR priority.
- Determine the hash mask length.
- Determine the IPv6 ACL rule defining a legal BSR address range.
- Determine the BS period.
- Determine the BS timeout.
- Determine the IPv6 ACL rule for register message filtering.
- Determine the register suppression time.
- Determine the register probe time.
- Determine the IPv6 ACL rule and sequencing rule for disabling an SPT switchover.

Enabling IPv6 PIM-SM

With IPv6 PIM-SM enabled, a router sends hello messages periodically to discover IPv6 PIM neighbors and processes messages from the IPv6 PIM neighbors. When you deploy an IPv6 PIM-SM domain, enable IPv6 PIM-SM on all non-border interfaces of the routers.

ⓘ IMPORTANT:

All the interfaces of the same device must operate in the same IPv6 PIM mode.

To enable IPv6 PIM-SM:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IPv6 multicast routing.	multicast ipv6 routing-enable	Disabled by default.
3. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Enable IPv6 PIM-SM.	pim ipv6 sm	Disabled by default.

For more information about the **multicast ipv6 routing-enable** command, see *IP Multicast Command Reference*.

Configuring an RP

An RP can be manually configured or dynamically elected through the BSR mechanism. For a large IPv6 PIM network, static RP configuration is a tedious job. Generally, static RP configuration is just a backup

method for the dynamic RP election mechanism to enhance the robustness and operation manageability of a multicast network.

! **IMPORTANT:**

In an IPv6 PIM network, if both IPv6 PIM-SM and IPv6 BIDIR-PIM are enabled, do not configure the same RP to serve IPv6 PIM-SM and IPv6 BIDIR-PIM simultaneously to avoid IPv6 PIM routing table errors.

Configuring a static RP

If only one dynamic RP exists in a network, manually configuring a static RP can avoid communication interruption because of single-point failures. It can also avoid frequent message exchange between C-RPs and the BSR.

! **IMPORTANT:**

To enable a static RP to work normally, you must perform this configuration on all routers in the IPv6 PIM-SM domain and specify the same RP address.

Perform the following configuration on all the routers in the IPv6 PIM-SM domain.

To configure a static RP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Configure a static RP for IPv6 PIM-SM.	static-rp <i>ipv6-rp-address</i> [<i>acl6-number</i>] [preferred]	No static RP by default.

Configuring a C-RP

In an IPv6 PIM-SM domain, you can configure routers that intend to become the RP as C-RPs. The BSR collects the C-RP information by receiving the C-RP-Adv messages from C-RPs or auto-RP announcements from other routers and organizes the information into an RP-set, which is flooded throughout the entire network. Then, the other routers in the network calculate the mappings between specific group ranges and the corresponding RPs based on the RP-set. HP recommends that you configure C-RPs on backbone routers.

To guard against C-RP spoofing, you need to configure a legal C-RP address range and the range of IPv6 multicast groups to be served on the BSR. In addition, because every C-BSR has a chance to become the BSR, you need to configure the same filtering policy on all C-BSRs in the IPv6 PIM-SM domain.

When you configure a C-RP, ensure a relatively large bandwidth between this C-RP and the other devices in the IPv6 PIM-SM domain.

An RP can serve multiple IPv6 multicast groups or all IPv6 multicast groups. Only one RP can forward IPv6 multicast traffic for an IPv6 multicast group at a moment.

To configure a C-RP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A

Step	Command	Remarks
3. Configure an interface to be a C-RP for IPv6 PIM-SM.	c-rp <i>ipv6-address</i> [{ group-policy <i>acl6-number</i> scope <i>scope-id</i> } priority <i>priority</i> holdtime <i>hold-interval</i> advertisement-interval <i>adv-interval</i>] *	No C-RPs are configured by default.
4. Configure a legal C-RP address range and the range of IPv6 multicast groups to be served.	crp-policy <i>acl6-number</i>	Optional. No restrictions by default.

Enabling embedded RP

With the embedded RP feature enabled, the router can resolve the RP address directly from the IPv6 multicast group address of an IPv6 multicast packets. This RP can replace the statically configured RP or the RP dynamically calculated based on the BSR mechanism. Therefore, the DR does not need to identify the RP address beforehand.

Perform this configuration on all routers in the IPv6 PIM-SM domain.

To enable embedded RP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Enable embedded RP.	embedded-rp [<i>acl6-number</i>]	Optional. By default, embedded RP is enabled for IPv6 multicast groups in the default embedded RP address scopes.

NOTE:

The default embedded RP address scopes are FF7x::/12 and FFFx::/12. Here "x" refers to any legal address scope. For more information about the Scope field, see "[Multicast overview](#)."

Configuring C-RP timers globally

To enable the BSR to distribute the RP-set information within the IPv6 PIM-SM domain, C-RPs must periodically send C-RP-Adv messages to the BSR. The BSR obtains the RP-set information from the received messages, and encapsulates its own IPv6 address together with the RP-set information in its bootstrap messages. The BSR then floods the bootstrap messages to all IPv6 routers in the network.

Each C-RP encapsulates a timeout value in its C-RP-Adv messages. After receiving a C-RP-Adv message, the BSR obtains this timeout value and starts a C-RP timeout timer. If the BSR fails to obtain a subsequent C-RP-Adv message from the C-RP when the timer times out, the BSR assumes the C-RP to have expired or become unreachable.

You must configure the C-RP timers on C-RP routers.

To configure C-RP timers globally:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A

Step	Command	Remarks
3. Configure the C-RP-Adv interval.	c-rp advertisement-interval <i>interval</i>	Optional. 60 seconds by default.
4. Configure C-RP timeout time.	c-rp holdtime <i>interval</i>	Optional. 150 seconds by default.

For more information about the configuration of other timers in IPv6 PIM-SM, see "[Configuring IPv6 PIM common timers.](#)"

Configuring a BSR

Configuration guidelines

Before you configure a BSR, make sure that you are familiar with BSR election process, BSR legal address against BSR spoofing, IPv6 PIM domain border, global C-BSR parameters, C-BSR timers, and bootstrap message fragments (BSMFs).

- BSR election process

An IPv6 PIM-SM domain can have only one BSR, but must have at least one C-BSR. Any router can be configured as a C-BSR. Elected from C-BSRs, the BSR is responsible for collecting and advertising RP information in the IPv6 PIM-SM domain.

You should configure C-BSRs on routers in the backbone network. When you configure a router as a C-BSR, be sure to specify the IPv6 address of an IPv6 PIM-SM-enabled interface on the router. The BSR election process is as follows:

- a. Initially, every C-BSR assumes itself to be the BSR of this IPv6 PIM-SM domain and uses its interface IPv6 address as the BSR address to send bootstrap messages.
 - b. When a C-BSR receives the bootstrap message of another C-BSR, it first compares its own priority with the other C-BSR's priority carried in the message. The C-BSR with a higher priority wins. If a tie exists in the priority, the C-BSR with a higher IPv6 address wins. The loser uses the winner's BSR address to replace its own BSR address and no longer assumes itself to be the BSR, and the winner keeps its own BSR address and continues assuming itself to be the BSR.
- BSR legal address against BSR spoofing

Configuring a legal range of BSR addresses enables filtering of bootstrap messages based on the address range, thereby preventing a maliciously configured host from masquerading as a BSR. You must make the same configuration on all routers in the IPv6 PIM-SM domain. Typical BSR spoofing cases and the corresponding preventive measures are as follows:

- Some maliciously configured hosts can forge bootstrap messages to fool routers and change RP mappings. Such attacks often occur on border routers. Because a BSR is inside the network whereas hosts are outside the network, you can protect a BSR against attacks from external hosts by enabling the border routers to perform neighbor checks and RPF checks on bootstrap messages and to discard unwanted messages.
- If an attacker controls a router in the network or if the network contains an illegal router, the attacker can configure this router as a C-BSR and make it win BSR election to control the right of advertising RP information in the network. After you configure a router as a C-BSR, the router automatically floods the network with bootstrap messages. Because a bootstrap message has a hop limit value of 1, the whole network will not be affected as long as the neighbor router discards these bootstrap messages. Therefore, with a legal BSR address range configured on

all routers in the entire network, all these routers will discard bootstrap messages from out of the legal address range.

These preventive measures can partially protect the security of BSRs in a network. However, if an attacker controls a legal BSR, the problem will still occur.

Because a large amount of information needs to be exchanged between a BSR and the other devices in the IPv6 PIM-SM domain, a relatively large bandwidth should be provided between the C-BSR and the other devices in the IPv6 PIM-SM domain.

- IPv6 PIM domain border

As the administrative core of an IPv6 PIM-SM domain, the BSR sends the collected RP-set information in the form of bootstrap messages to all routers in the IPv6 PIM-SM domain.

An IPv6 PIM domain border is a bootstrap message boundary. Each BSR has its specific service scope. IPv6 PIM domain border interfaces partition a network into different IPv6 PIM-SM domains. Bootstrap messages cannot cross a domain border in either direction.

- C-BSR parameters

In each IPv6 PIM-SM domain, a unique BSR is elected from C-BSRs. The C-RPs in the IPv6 PIM-SM domain send advertisement messages to the BSR. The BSR summarizes the advertisement messages to form an RP-set and advertises it to all routers in the IPv6 PIM-SM domain. All the routers use the same hash algorithm to get the RP address that corresponds to specific IPv6 multicast groups.

- C-BSR timers

The BSR election winner multicasts its own IPv6 address and RP-Set information throughout the region that it serves through bootstrap messages. The BSR floods bootstrap messages throughout the network at the interval of the BS (BSR state) period. Any C-BSR that receives a bootstrap message retains the RP-set for the length of BS timeout, during which no BSR election occurs. If no bootstrap message is received from the BSR before the BS timeout timer expires, a new BSR election process begins among the C-BSRs.

About the BS period:

- By default, the BS period is determined by the formula "BS period = (BS timeout – 10) / 2." The default BS timeout is 130 seconds, so the default BS period = (130 – 10) / 2 = 60 (seconds).
- If this parameter is manually configured, the system will use the configured value.

About the BS timeout timer:

- By default, the BS timeout value is determined by the formula "BS timeout timer = BS period × 2 + 10." The default BS period is 60 seconds, so the default BS timeout timer = 60 × 2 + 10 = 130 (seconds).
- If this parameter is manually configured, the system will use the configured value.

In the configuration, make sure that the BS period is smaller than the BS timeout value.

- Bootstrap message fragments (BSMFs)

Generally, a BSR periodically distributes the RP-set information in bootstrap messages within the IPv6 PIM-SM domain. It encapsulates a BSM in an IPv6 datagram and might split the datagram into fragments if the message exceeds the maximum transmission unit (MTU). In respect of such IP fragmentation, loss of a single IP fragment leads to unavailability of the entire message.

Semantic fragmentation of BSMs can solve this issue. When a BSM exceeds the MTU, it is split to multiple bootstrap message fragments (BSMFs).

- After receiving a BSMF that contains the RP-set information of one group range, a non-BSR router updates corresponding RP-set information directly.

- If the RP-set information of one group range is carried in multiple BSMFs, a non-BSR router updates corresponding RP-set information after receiving all these BSMFs.

Because the RP-set information contained in each segment is different, loss of some IP fragments will not result in dropping of the entire message.

The function of BSM semantic fragmentation is enabled by default. Devices not supporting this function might deem a fragment as an entire message and learn only part of the RP-set information. Therefore, if such devices exist in the IPv6 PIM-SM domain, you need to disable the semantic fragmentation function on the C-BSRs.

Generally, a BSR performs BSM semantic fragmentation according to the MTU of its BSR interface. However, the semantic fragmentation of BSMs originated because of learning of a new IPv6 PIM neighbor is performed according to the MTU of the outgoing interface.

Configuring a C-BSR

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Configure an interface as a C-BSR.	c-bsr <i>ipv6-address</i> [<i>hash-length</i> [<i>priority</i>]]	No C-BSRs are configured by default.
4. Configure a legal BSR address range.	bsr-policy <i>acl6-number</i>	Optional. No restrictions by default.

Configuring an IPv6 PIM domain border

Perform the following configuration on routers that you want to configure as an IPv6 PIM domain border.

To configure an IPv6 PIM border domain:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure an IPv6 PIM domain border.	pim ipv6 bsr-boundary	No IPv6 PIM domain border is configured by default.

Configuring C-BSR parameters globally

Perform the following configuration on C-BSR routers.

To configure C-BSR parameters globally:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Configure the hash mask length.	c-bsr hash-length <i>hash-length</i>	Optional. 126 by default.

Step	Command	Remarks
4. Configure the C-BSR priority.	c-bsr priority <i>priority</i>	Optional. 64 by default.

Configuring C-BSR timers

Perform the following configuration on C-BSR routers.

To configure C-BSR timers:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Configure the BS period.	c-bsr interval <i>interval</i>	Optional. For the default value, see the note after this table.
4. Configure the BS timeout timer.	c-bsr holdtime <i>interval</i>	Optional. For the default value, see the note after this table.

Disabling BSM semantic fragmentation

To disable the BSM semantic fragmentation function:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Disable the BSM semantic fragmentation function.	undo bsm-fragment enable	By default, the BSM semantic fragmentation function is enabled.

Configuring IPv6 administrative scoping

With IPv6 administrative scoping disabled, an IPv6 PIM-SM domain has only one BSR. The BSR manages the whole network. To manage your network more effectively and specifically, you can partition the IPv6 PIM-SM domain into multiple IPv6 admin-scope zones. Each IPv6 admin-scope zone maintains a BSR, which serves a specific IPv6 multicast group range. The IPv6 global scope zone also maintains a BSR, which serves the IPv6 multicast groups with the Scope field in the group addresses being 14.

Enabling IPv6 administrative scoping

Before you configure an IPv6 admin-scope zone, you must enable IPv6 administrative scoping.

Perform the following configuration on all routers in the IPv6 PIM-SM domain.

To enable IPv6 administrative scoping:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A

Step	Command	Remarks
3.	Enable IPv6 administrative scoping. c-bsr admin-scope	Disabled by default

Configuring an IPv6 admin-scope zone boundary

The boundary of each IPv6 admin-scope zone is formed by ZBRs. Each admin-scope zone maintains a BSR, which serves multicast groups with a specific Scope field in their group addresses. Multicast protocol packets (such as assert messages and bootstrap messages) that belong to this range cannot cross the admin-scope zone boundary.

Perform the following configuration on routers that you want to configure as a ZBR.

To configure an IPv6 admin-scope zone boundary:

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Enter interface view. interface interface-type interface-number	N/A
3.	Configure an IPv6 multicast forwarding boundary. multicast ipv6 boundary { <i>ipv6-group-address prefix-length</i> scope { <i>scope-id</i> admin-local global organization-local site-local } }	By default, no multicast forwarding boundary is configured.

For more information about the **multicast ipv6 boundary** command, see *IP Multicast Command Reference*.

Configuring C-BSRs for IPv6 admin-scope zones

In a network with IPv6 administrative scoping enabled, BSRs are elected from C-BSRs specific to different Scope field values. The C-RPs in the network send advertisement messages to the specific BSR. The BSR summarizes the advertisement messages to form an RP-set and advertises it to all routers in the specific admin-scope zone. All the routers use the same hash algorithm to get the RP address corresponding to the specific IPv6 multicast group.

You can configure the hash mask length and C-BSR priority globally and in an IPv6 admin-scope zone.

- The values configured in the IPv6 admin-scope zone have preference over the global values.
- If you do not configure these parameters in the IPv6 admin-scope zone, the corresponding global values will be used.

For configuration of global C-BSR parameters, see "[Configuring C-BSR parameters globally.](#)"

Perform the following configuration on the routers that you want to configure as C-BSRs in IPv6 admin-scope zones.

To configure a C-BSR for an IPv6 admin-scope zone:

Step	Command	Remarks
1.	Enter system view. system-view	N/A
2.	Enter IPv6 PIM view. pim ipv6	N/A
3.	Configure a C-BSR for an IPv6 admin-scope zone. c-bsr scope { <i>scope-id</i> admin-local global organization-local site-local } [hash-length <i>hash-length</i> priority <i>priority</i>] *	No C-BSRs are configured for an IPv6 admin-scope zone by default.

Configuring IPv6 multicast source registration

Configuration guidelines

Within an IPv6 PIM-SM domain, the source-side DR sends register messages to the RP, and these register messages have different IPv6 multicast source or IPv6 multicast group addresses. You can configure a filtering rule to filter register messages so that the RP can serve specific IPv6 multicast groups. If the filtering rule denies an (S, G) entry, or if the filtering rule does not define an action for this entry, the RP will send a register-stop message to the DR to stop the registration process for the IPv6 multicast data.

In view of information integrity of register messages in the transmission process, you can configure the device to calculate the checksum based on the entire register messages. However, to reduce the workload of encapsulating data in register messages and for the sake of interoperability, HP does not recommend this method of checksum calculation.

When receivers stop receiving data addressed to a certain IPv6 multicast group through the RP (that is, the RP stops serving the receivers of that IPv6 multicast group), or when the RP starts receiving IPv6 multicast data from the IPv6 multicast source along the SPT, the RP sends a register-stop message to the source-side DR. After receiving this message, the DR stops sending register messages encapsulated with IPv6 multicast data and starts a register-stop timer. Before the register-stop timer expires, the DR sends a null register message (a register message without multicast data) to the RP. If the DR receives a register-stop message during the register probe time, it will reset its register-stop timer. Otherwise, the DR starts sending register messages with encapsulated data again when the register-stop timer expires.

The register-stop timer is set to a random value chosen uniformly from the interval (0.5 times register_suppression_time, 1.5 times register_suppression_time) minus register_probe_time.

Configure a filtering rule for register messages on all C-RP routers, and configure them to calculate the checksum based on the entire register messages. Configure the register suppression time and the register probe time on all routers that might become IPv6 source-side DRs.

Configuration procedure

To configure register-related parameters:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Configure a filtering rule for register messages.	register-policy <i>acl6-number</i>	Optional. No register filtering rule by default.
4. Configure the device to calculate the checksum based on the entire register messages.	register-whole-checksum	Optional. Based on the header of register messages by default.
5. Configure the register suppression time.	register-suppression-timeout <i>interval</i>	Optional. 60 seconds by default.
6. Configure the register probe time.	probe-interval <i>interval</i>	Optional. 5 seconds by default.

Disabling SPT switchover

If an HP 5500 EI switch acts as an RP or the receiver-side DR, it initiates an SPT switchover process by default upon receiving the first IPv6 multicast packet along the RPT. You can disable the switchover from RPT to SPT.

! **IMPORTANT:**

For an HP 5500 EI switch, once an IPv6 multicast forwarding entry is created, subsequent IPv6 multicast data will not be encapsulated in register messages before being forwarded even if a register outgoing interface is available. Therefore, to avoid forwarding failure, do not use the **spt-switch-threshold infinity** command on a switch that might become an RP (namely, a static RP or a C-RP).

To configure SPT switchover:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Disable the SPT switchover.	spt-switch-threshold infinity [group-policy <i>acl6-number</i> [order <i>order-value</i>]]	Optional. By default, the device switches to the SPT immediately after it receives the first IPv6 multicast packet from the RPT.

Configuring IPv6 BIDIR-PIM

IPv6 BIDIR-PIM configuration task list

Task	Remarks	
Enabling IPv6 PIM-SM	Required.	
Enabling IPv6 BIDIR-PIM	Required.	
Configuring an RP	Configuring a static RP	Required.
	Configuring a C-RP	Use any approach.
	Enabling embedded RP	
	Configuring C-RP timers globally	Optional.
Configuring a BSR	Configuring a C-BSR	Required.
	Configuring an IPv6 BIDIR-PIM domain border	Optional.
	Configuring global C-BSR parameters	Optional.
	Configuring C-BSR timers	Optional.
	Disabling BSM semantic fragmentation	Optional.
Configuring IPv6 administrative scoping	Enabling IPv6 administrative scoping	Optional.
	Configuring an IPv6 admin-scope zone boundary	Optional.
	Configuring C-BSRs for each admin-scope zone	Optional.

Task	Remarks
Configuring IPv6 PIM common features	Optional.

Configuration prerequisites

Before you configure IPv6 BIDIR-PIM, complete the following tasks:

- Enable IPv6 forwarding and configure an IPv6 unicast routing protocol so that all devices in the domain can communicate with each other at Layer 3.
- Determine the IPv6 address of a static RP and the IPv6 ACL that defines the range of IPv6 multicast groups to be served by the static RP.
- Determine the C-RP priority and the IPv6 ACL that defines the range of IPv6 multicast groups to be served by each C-RP.
- Determine the legal C-RP address range and the IPv6 ACL that defines the range of IPv6 multicast groups to be served.
- Determine the C-RP-Adv interval.
- Determine the C-RP timeout.
- Determine the C-BSR priority.
- Determine the hash mask length.
- Determine the IPv6 ACL defining the legal BSR address range.
- Determine the BS period.
- Determine the BS timeout.

Enabling IPv6 PIM-SM

You must enable IPv6 PIM-SM before enabling IPv6 BIDIR-PIM because IPv6 BIDIR-PIM is implemented on the basis of IPv6 PIM-SM. To deploy an IPv6 BIDIR-PIM domain, enable IPv6 PIM-SM on all non-border interfaces of the domain.

ⓘ IMPORTANT:

On a router, all interfaces in the same VPN instance must operate in the same IPv6 PIM mode.

To enable IPv6 PIM-SM:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IPv6 multicast routing.	multicast ipv6 routing-enable	Disabled by default
3. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Enable IPv6 PIM-SM.	pim ipv6 sm	Disabled by default

For more information about the **multicast ipv6 routing-enable** command, see *IP Multicast Command Reference*.

Enabling IPv6 BIDIR-PIM

Perform this configuration on all routers in the IPv6 BIDIR-PIM domain.

To enable IPv6 BIDIR-PIM:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Enable IPv6 BIDIR-PIM.	bidir-pim enable	Disabled by default

Configuring an RP

An RP can be manually configured or dynamically elected through the BSR mechanism. For a large IPv6 PIM network, static RP configuration is a tedious job. Generally, static RP configuration is just a backup means for the dynamic RP election mechanism to enhance the robustness and operation manageability of a multicast network.

! IMPORTANT:

In an IPv6 PIM network, if both IPv6 PIM-SM and IPv6 BIDIR-PIM are enabled, do not configure the same RP to serve IPv6 PIM-SM and IPv6 BIDIR-PIM simultaneously to avoid IPv6 PIM routing table errors.

Configuring a static RP

If only one dynamic RP exists in a network, manually configuring a static RP can avoid communication interruption because of single-point failures and avoid frequent message exchange between C-RPs and the BSR.

In IPv6 BIDIR-PIM, a static RP can be specified with a virtual IPv6 address. For example, if the IPv6 addresses of the interfaces at the two ends of a link are 1001::1/64 and 1001::2/64, you can specify a virtual IPv6 address, like 1001::100/64, for the static RP. As a result, the link becomes an RPL.

! IMPORTANT:

You must perform static RP configuration on all routers in the IPv6 PIM-SM domain and specify the same RP address.

Perform this configuration on all routers in the IPv6 BIDIR-PIM domain.

To configure a static RP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Configure a static RP for IPv6 BIDIR-PIM.	static-rp <i>ipv6-rp-address</i> [<i>acl6-number</i>] [preferred] bidir	No static RP by default

Configuring a C-RP

In an IPv6 BIDIR-PIM domain, you can configure routers that intend to become the RP as C-RPs. The BSR collects the C-RP information by receiving the C-RP-Adv messages from C-RPs or auto-RP announcements

from other routers and organizes the information into an RP-set, which is flooded throughout the entire network. Then, the other routers in the network calculate the mappings between specific group ranges and the corresponding RPs based on the RP-set. HP recommends that you configure C-RPs on backbone routers.

To guard against C-RP spoofing, configure a legal C-RP address range and the range of multicast groups to be served on the BSR. In addition, because every C-BSR has a chance to become the BSR, you must configure the same filtering policy on all C-BSRs in the IPv6 BIDIR-PIM domain.

To configure a C-RP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Configure an interface to be a C-RP for IPv6 BIDIR-PIM.	c-rp <i>ipv6-address</i> [{ group-policy <i>acl6-number</i> scope <i>scope-id</i> } priority <i>priority</i> holdtime <i>hold-interval</i> advertisement-interval <i>adv-interval</i>] * bidir	No C-RP is configured by default.

NOTE:

- When you configure a C-RP, ensure a relatively large bandwidth between this C-RP and the other devices in the IPv6 BIDIR-PIM domain.
- An RP can serve multiple IPv6 multicast groups or all IPv6 multicast groups. Only one RP can forward multicast traffic for an IPv6 multicast group at a moment.

Enabling embedded RP

With the embedded RP feature enabled, the router can resolve the RP address directly from the IPv6 multicast group address of an IPv6 multicast packets. This RP can replace the statically configured RP or the RP dynamically calculated based on the BSR mechanism. Thus, the DR does not need to know the RP address beforehand.

Perform this configuration on all routers in the IPv6 BIDIR-PIM domain.

To enable embedded RP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Enable embedded RP.	embedded-rp [<i>acl6-number</i>]	Optional. By default, embedded RP is enabled for IPv6 multicast groups in the default embedded RP address scopes.

NOTE:

The default embedded RP address scopes are FF7x::/12 and FFFx::/12, where x refers to any legal address scope. For more information about the Scope field, see "[Multicast overview](#)."

Configuring C-RP timers globally

To enable the BSR to distribute the RP-Set information within the IPv6 BIDIR-PIM domain, C-RPs must periodically send C-RP-Adv messages to the BSR. The BSR learns the RP-Set information from the received messages, and encapsulates its own IPv6 address together with the RP-Set information in its bootstrap messages. The BSR then floods the bootstrap messages to all IPv6 routers in the network.

Each C-RP encapsulates a timeout value in its C-RP-Adv messages. After receiving a C-RP-Adv message, the BSR obtains this timeout value and starts a C-RP timeout timer. If the BSR fails to hear a subsequent C-RP-Adv message from the C-RP when the timer times out, the BSR assumes the C-RP to have expired or become unreachable.

The C-RP timers need to be configured on C-RP routers.

To configure C-RP timers globally:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Configure the C-RP-Adv interval.	c-rp advertisement-interval <i>interval</i>	Optional. 60 seconds by default.
4. Configure C-RP timeout time.	c-rp holdtime <i>interval</i>	Optional. 150 seconds by default.

For more information about the configuration of other timers in IPv6 PIM-SM, see "[Configuring IPv6 PIM common timers.](#)"

Configuring a BSR

Configuration guidelines

Before you configure a BSR, make sure that you are familiar with BSR election process, BSR legal address against BSR spoofing, IPv6 PIM domain border, global C-BSR parameters, C-BSR timers, and bootstrap message fragments (BSMFs). They operate the same way as in the PIM-SM domain. For more information, see "[Configuration guidelines.](#)"

Configuring a C-BSR

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Configure an interface as a C-BSR.	c-bsr <i>ipv6-address</i> [<i>hash-length</i> [<i>priority</i>]]	No C-BSRs are configured by default.
4. Configure a legal BSR address range.	bsr-policy <i>acl6-number</i>	Optional. No restrictions on BSR address range by default.

Configuring an IPv6 BIDIR-PIM domain border

Perform the following configuration on routers that you want to configure as the IPv6 BIDIR-PIM domain border.

To configure an IPv6 BIDIR-PIM domain border:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure an IPv6 BIDIR-PIM domain border.	pim ipv6 bsr-boundary	By default, no IPv6 BIDIR-PIM domain border is configured.

Configuring global C-BSR parameters

Perform the following configuration on C-BSR routers.

To configure global C-BSR parameters:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Configure the hash mask length.	c-bsr hash-length <i>hash-length</i>	Optional. 126 by default.
4. Configure the C-BSR priority.	c-bsr priority <i>priority</i>	Optional. 64 by default.

Configuring C-BSR timers

Perform the following configuration on C-BSR routers.

To configure C-BSR timers:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Configure the BS period.	c-bsr interval <i>interval</i>	Optional. For the default value, see the note after this table.
4. Configure the BS timeout timer.	c-bsr holdtime <i>interval</i>	Optional. For the default value, see the note after this table.

Disabling BSM semantic fragmentation

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Disable the BSM semantic fragmentation function.	undo bsm-fragment enable	By default, the BSM semantic fragmentation function is enabled.

Configuring IPv6 administrative scoping

With administrative scoping disabled, an IPv6 BIDIR-PIM domain has only one BSR. The BSR manages the whole network. To manage your network more effectively and specifically, you can partition the IPv6 BIDIR-PIM domain into multiple admin-scope zones. Each admin-scope zone maintains a BSR, which serves a specific multicast group range. The global scope zone also maintains a BSR, which serves all the rest multicast groups.

Enabling IPv6 administrative scoping

Before you configure an IPv6 admin-scope zone, you must enable IPv6 administrative scoping first.

Perform the following configuration on all routers in the IPv6 PIM-SM domain.

To enable IPv6 administrative scoping:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Enable IPv6 administrative scoping.	c-bsr admin-scope	Disabled by default

Configuring an IPv6 admin-scope zone boundary

The boundary of each IPv6 admin-scope zone is formed by ZBRs. Each admin-scope zone maintains a BSR, which serves a specific IPv6 multicast group range. IPv6 multicast packets (such as assert messages and bootstrap messages) that belong to this range cannot cross the admin-scope zone boundary.

Perform the following configuration on routers that you want to configure as a ZBR.

To configure an admin-scope zone boundary:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure an IPv6 multicast forwarding boundary.	multicast ipv6 boundary { <i>ipv6-group-address prefix-length</i> scope { <i>scope-id</i> admin-local global organization-local site-local } }	By default, no IPv6 multicast forwarding boundary is configured.

For more information about the **multicast ipv6 boundary** command, see *IP Multicast Command Reference*.

Configuring C-BSRs for each admin-scope zone

In a network with administrative scoping enabled, group-range-specific BSRs are elected from C-BSRs. C-RPs in the network send advertisement messages to the specific BSR. The BSR summarizes the advertisement messages to form an RP-set and advertises it to all routers in the specific admin-scope zone. All the routers use the same hash algorithm to get the RP address corresponding to the specific multicast group.

You can configure the hash mask length and C-BSR priority globally, only in an IPv6 admin-scope zone, or both globally and in an IPv6 admin-scope zone.

- The values configured in the IPv6 admin-scope zone have preference over the global values.
- If you do not configure these parameters in the IPv6 admin-scope zone, the corresponding global values will be used.

For configuration of global C-BSR parameters, see "[Configuring global C-BSR parameters](#)."

Perform the following configuration on the routers that you want to configure as C-BSRs in admin-scope zones.

To configure a C-BSR for an admin-scope zone:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Configure a C-BSR for an admin-scope zone.	c-bsr scope { <i>scope-id</i> admin-local global organization-local site-local } [hash-length <i>hash-length</i> priority <i>priority</i>] *	No C-BSRs are configured for an admin-scope zone by default.

Configuring IPv6 PIM-SSM

ⓘ IMPORTANT:

The IPv6 PIM-SSM model needs the support of MLDv2. Be sure to enable MLDv2 on IPv6 PIM routers with receivers attached to them.

IPv6 PIM-SSM configuration task list

Task	Remarks
Enabling IPv6 PIM-SSM	Required
Configuring the IPv6 SSM group range	Optional
Configuring IPv6 PIM common features	Optional

Configuration prerequisites

Before you configure IPv6 PIM-SSM, complete the following tasks:

- Enable IPv6 forwarding and configure an IPv6 unicast routing protocol so that all devices in the domain are interoperable at the network layer.

- Determine the IPv6 SSM group range.

Enabling IPv6 PIM-SM

Configuration guidelines

The SSM model is implemented based on some subsets of IPv6 PIM-SM. Therefore, you must enable IPv6 PIM-SM before configuring IPv6 PIM-SSM.

When you deploy an IPv6 PIM-SSM domain, enable IPv6 PIM-SM on all non-border interfaces of routers.

All the interfaces of the same device must operate in the same IPv6 PIM mode.

Configuration procedure

To enable IPv6 PIM-SSM:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable IPv6 multicast routing.	multicast ipv6 routing-enable	Disabled by default
3. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Enable IPv6 PIM-SM.	pim ipv6 sm	Disabled by default

For more information about the **multicast ipv6 routing-enable** command, see *IP Multicast Command Reference*.

Configuring the IPv6 SSM group range

Configuration guidelines

Whether the information from an IPv6 multicast source is delivered to the receivers based on the IPv6 PIM-SSM model or the IPv6 PIM-SM model depends on whether the group address in the (S, G) channel subscribed by the receivers falls into the IPv6 SSM group range. All IPv6 PIM-SM-enabled interfaces assume that IPv6 multicast groups within this address range are using the IPv6 SSM model.

Make sure that the same IPv6 SSM group range is configured on all routers in the entire domain. Otherwise, IPv6 multicast data cannot be delivered through the IPv6 SSM model.

When a member of an IPv6 multicast group in the IPv6 SSM group range sends an MLDv1 report message, the device does not trigger a (*, G) join.

Configuration procedure

Perform the following configuration on all routers in the IPv6 PIM-SSM domain.

To configure the IPv6 SSM group range:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A

Step	Command	Remarks
3. Configure the IPv6 SSM group range.	<code>ssm-policy acl6-number</code>	Optional. FF3x::/32 by default, here "x" refers to any legal group scope.

Configuring IPv6 PIM common features

For the configuration tasks in this section:

- In IPv6 PIM view, the configuration is effective on all interfaces. In interface view, the configuration is effective on only the current interface.
- If the same function or parameter is configured in both IPv6 PIM view and interface view, the configuration in interface view has preference over the configuration in PIM view, regardless of the configuration sequence.

IPv6 PIM common feature configuration task list

Task	Remarks
Configuring an IPv6 multicast data filter	Optional
Configuring a hello message filter	Optional
Configuring IPv6 PIM hello options	Optional
Configuring the prune delay	Optional
Configuring IPv6 PIM common timers	Optional
Configuring join/prune message sizes	Optional
Configuring IPv6 PIM to work with BFD	Optional
Setting the DSCP value for IPv6 PIM messages	Optional

Configuration prerequisites

Before you configure IPv6 PIM common features, complete the following tasks:

- Enable IPv6 forwarding and configure an IPv6 unicast routing protocol so that all devices in the domain are interoperable at the network layer.
- Configure IPv6 PIM-DM (or IPv6 PIM-SM or IPv6 PIM-SSM).
- Determine the IPv6 ACL rule for filtering IPv6 multicast data.
- Determine the IPv6 ACL rule defining a legal source address range for hello messages.
- Determine the priority for DR election (global value/interface level value).
- Determine the IPv6 PIM neighbor timeout time (global value/interface value).
- Determine the prune message delay (global value/interface level value).
- Determine the prune override interval (global value/interface level value).
- Determine the prune delay.
- Determine the hello interval (global value/interface level value).
- Determine the maximum delay between hello message (interface level value).

- Determine the assert timeout time (global value/interface value).
- Determine the join/prune interval (global value/interface level value).
- Determine the join/prune timeout (global value/interface value).
- Determine the IPv6 multicast source lifetime.
- Determine the maximum size of join/prune messages.
- Determine the maximum number of (S, G) entries in a join/prune message.
- Determine the DSCP value for IPv6 PIM messages.

Configuring an IPv6 multicast data filter

Configuration guidelines

In either an IPv6 PIM-DM domain or an IPv6 PIM-SM domain, routers can check passing-by IPv6 multicast data based on the configured filtering rules and determine whether to continue forwarding the IPv6 multicast data. In other words, IPv6 PIM routers can act as IPv6 multicast data filters. These filters can help implement traffic control on one hand, and control the information available to downstream receivers to enhance data security on the other hand.

Generally, a smaller distance from the filter to the IPv6 multicast source results in a more remarkable filtering effect.

This filter works not only on independent IPv6 multicast data but also on IPv6 multicast data encapsulated in register messages.

Configuration procedure

To configure an IPv6 multicast data filter:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Configure an IPv6 multicast group filter.	source-policy <i>acl6-number</i>	No IPv6 multicast data filter by default

Configuring a hello message filter

Along with the wide applications of IPv6 PIM, the security requirement for the protocol is becoming increasingly demanding. The establishment of correct IPv6 PIM neighboring relationships is a prerequisite for secure application of IPv6 PIM. To guide against IPv6 PIM message attacks, you can configure a legal source address range for hello messages on interfaces of routers to ensure the correct IPv6 PIM neighboring relationships.

To configure a hello message filter:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A

Step	Command	Remarks
3. Configure a hello message filter.	<code>pim ipv6 neighbor-policy acl6-number</code>	No hello message filter by default

NOTE:

With the hello message filter configured, if hello messages of an existing IPv6 PIM neighbor fail to pass the filter, the IPv6 PIM neighbor will be removed automatically when it times out.

Configuring IPv6 PIM hello options

IPv6 PIM hello options

In either an IPv6 PIM-DM domain or an IPv6 PIM-SM domain, the hello messages sent among routers contain the following configurable options:

- **DR_Priority** (for IPv6 PIM-SM only)—Priority for DR election. The higher the priority is, the easier it is for the router to win DR election. You can configure this parameter on all the routers in a multi-access network directly connected to IPv6 multicast sources or receivers.
- **Holdtime**—The timeout time of IPv6 PIM neighbor reachability state. When this timer times out, if the router has received no hello message from an IPv6 PIM neighbor, it assumes that this neighbor has expired or become unreachable.
- **LAN_Prune_Delay**—The delay of prune messages on a multi-access network. This option consists of LAN-delay (namely, prune message delay), override-interval, and neighbor tracking flag. If the LAN-delay or override-interval values of different IPv6 PIM routers on a multi-access subnet are different, the largest value takes effect. If you want to enable neighbor tracking, be sure to enable the neighbor tracking feature on all IPv6 PIM routers on a multi-access subnet.

The LAN-delay setting will cause the upstream routers to delay forwarding received prune messages. The override-interval sets the length of time that a downstream router can wait before sending a prune override message. When a router receives a prune message from a downstream router, it does not perform the prune action immediately. Instead, it maintains the current forwarding state for a period of LAN-delay plus override-interval. If the downstream router needs to continue receiving IPv6 multicast data, it must send a join message within the prune override interval. Otherwise, the upstream route will perform the prune action when the period of LAN-delay plus override-interval times out.

A hello message sent from an IPv6 PIM router contains a generation ID option. The generation ID is a random value for the interface on which the hello message is sent. Normally, the generation ID of an IPv6 PIM router does not change unless the status of the router changes (for example, when IPv6 PIM is just enabled on the interface or the device is restarted). When the router starts or restarts sending hello messages, it generates a new generation ID. If an IPv6 PIM router finds that the generation ID in a hello message from the upstream router has changed, it assumes that the status of the upstream neighbor is lost or that the upstream neighbor has changed. In this case, it triggers a join message for state update.

If you disable join suppression (namely, enable neighbor tracking), be sure to disable the join suppression feature on all IPv6 PIM routers on a multi-access subnet. Otherwise, the upstream router will fail to explicitly track join messages from downstream routers.

Configuring hello options globally

Step	Command	Remarks
1. Enter system view.	<code>system-view</code>	N/A

Step	Command	Remarks
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Configure the priority for DR election.	hello-option dr-priority <i>priority</i>	Optional. 1 by default.
4. Configure IPv6 PIM neighbor timeout time.	hello-option holdtime <i>interval</i>	Optional. 105 seconds by default.
5. Configure the prune message delay time (LAN-delay).	hello-option lan-delay <i>interval</i>	Optional. 500 milliseconds by default.
6. Configure the prune override interval.	hello-option override-interval <i>interval</i>	Optional. 2,500 milliseconds by default.
7. Disable join suppression.	hello-option neighbor-tracking	Enabled by default.

Configuring hello options on an interface

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the priority for DR election.	pim ipv6 hello-option dr-priority <i>priority</i>	Optional. 1 by default.
4. Configure IPv6 PIM neighbor timeout time.	pim ipv6 hello-option holdtime <i>interval</i>	Optional. 105 seconds by default.
5. Configure the prune message delay time (LAN-delay).	pim ipv6 hello-option lan-delay <i>interval</i>	Optional. 500 milliseconds by default.
6. Configure the prune override interval.	pim ipv6 hello-option override-interval <i>interval</i>	Optional. 2500 milliseconds by default.
7. Disable join suppression.	pim ipv6 hello-option neighbor-tracking	Enabled by default.
8. Configure the interface to reject hello messages without a generation ID.	pim ipv6 require-genid	By default, hello messages without Generation_ID are accepted.

Configuring the prune delay

Configuring the prune delay interval on an upstream router in a shared network segment can make the upstream router not perform the prune action immediately after receiving the prune message from its downstream router. Instead, the upstream router maintains the current forwarding state for a period of time that the prune delay interval defines. In this period, if the upstream router receives a join message from the downstream router, it cancels the prune action. Otherwise, it performs the prune action.

To configure the prune delay time

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Configure the prune delay interval.	prune delay <i>interval</i>	Optional. 3 seconds by default, which equals the prune pending time.

Configuring IPv6 PIM common timers

IPv6 PIM common timers

IPv6 PIM routers discover IPv6 PIM neighbors and maintain IPv6 PIM neighboring relationships with other routers by periodically sending hello messages.

After receiving a hello message, an IPv6 PIM router waits a random period, which is smaller than the maximum delay between hello messages, before sending a hello message. This avoids collisions that occur when multiple IPv6 PIM routers send hello messages simultaneously.

An IPv6 PIM router periodically sends join/prune messages to its upstream for state update. A join/prune message contains the join/prune timeout time. The upstream router sets a join/prune timeout timer for each pruned downstream interface.

Any router that has lost assert election will prune its downstream interface and maintain the assert state for a period of time. When the assert state times out, the assert loser will resume IPv6 multicast forwarding.

When a router fails to receive subsequent IPv6 multicast data from the IPv6 multicast source S, the router does not immediately delete the corresponding (S, G) entry. Instead, it maintains the (S, G) entry for a period of time—namely, the IPv6 multicast source lifetime—before deleting the (S, G) entry.

Configuring IPv6 PIM common timers globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Configure the hello interval.	timer hello <i>interval</i>	Optional. 30 seconds by default.
4. Configure the join/prune interval.	timer join-prune <i>interval</i>	Optional. 60 seconds by default.
5. Configure the join/prune timeout time.	holdtime join-prune <i>interval</i>	Optional. 210 seconds by default.
6. Configure assert timeout time.	holdtime assert <i>interval</i>	Optional. 180 seconds by default.
7. Configure the IPv6 multicast source lifetime.	source-lifetime <i>interval</i>	Optional. 210 seconds by default.

Configuring IPv6 PIM common timers on an interface

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the hello interval.	pim ipv6 timer hello <i>interval</i>	Optional. 30 seconds by default.
4. Configure the maximum delay between hello messages.	pim ipv6 triggered-hello-delay <i>interval</i>	Optional. 5 seconds by default.
5. Configure the join/prune interval.	pim ipv6 timer join-prune <i>interval</i>	Optional. 60 seconds by default.
6. Configure the join/prune timeout time.	pim ipv6 holdtime join-prune <i>interval</i>	Optional. 210 seconds by default.
7. Configure assert timeout time.	pim ipv6 holdtime assert <i>interval</i>	Optional 180 seconds by default.

NOTE:

If no special networking requirements are raised, use the default settings.

Configuring join/prune message sizes

A large size of a join/prune message might result in loss of a larger amount of information if a message is lost. You can set a small value for the size of each join/prune message to reduce the impact in case of the loss of a message.

By controlling the maximum number of (S, G) entries in a join/prune message, you can effectively reduce the number of (S, G) entries sent per unit of time.

! IMPORTANT:

If IPv6 PIM snooping-enabled switches are deployed in the IPv6 PIM network, be sure to set a value no greater than the IPv6 path MTU for the maximum size of each join/prune message on the receiver-side edge IPv6 PIM devices.

To configure join/prune message sizes:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Configure the maximum size of each join/prune message.	jp-pkt-size <i>packet-size</i>	Optional. 8100 bytes by default.
4. Configure the maximum number of (S, G) entries in a join/prune message.	jp-queue-size <i>queue-size</i>	Optional. 1020 by default.

Configuring IPv6 PIM to work with BFD

IPv6 PIM uses hello messages to elect a DR for a multi-access network. The elected DR will be the only multicast forwarder on the multi-access network.

If the DR fails, a new DR election process will start after the DR is aged out. However, it might take a long period of time. To start a new DR election process immediately after the original DR fails, you can enable IPv6 PIM to work with Bidirectional Forwarding Detection (BFD) on a multi-access network to detect failures of the links among IPv6 PIM neighbors. You must enable IPv6 PIM to work with BFD on all IPv6 PIM-capable routers on a multi-access network, so that the IPv6 PIM neighbors can fast detect DR failures and start a new DR election process.

Before you configure this feature on an interface, be sure to enable IPv6 PIM-DM or IPv6 PIM-SM on the interface.

To enable IPv6 PIM to work with BFD:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type interface-number</i>	N/A
3. Enable IPv6 PIM to work with BFD.	pim ipv6 bfd enable	Disabled by default

For more information about BFD, see *High Availability Configuration Guide*.

Setting the DSCP value for IPv6 PIM messages

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter IPv6 PIM view.	pim ipv6	N/A
3. Set the DSCP value for IPv6 PIM messages.	dscp <i>dscp-value</i>	Optional. By default, the DSCP value in IPv6 PIM messages is 48.

Displaying and maintaining IPv6 PIM

Task	Command	Remarks
Display the BSR information in the IPv6 PIM-SM domain and locally configured C-RP information in effect.	display pim ipv6 bsr-info [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about IPv6 unicast routes used by IPv6 PIM.	display pim ipv6 claimed-route [<i>ipv6-source-address</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display the number of IPv6 PIM control messages.	display pim ipv6 control-message counters [message-type { probe register register-stop }] [interface <i>interface-type interface-number</i> message-type { assert bsr crp graft graft-ack hello join-prune state-refresh }] * [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the DF information of IPv6 BIDIR-PIM.	display pim ipv6 df-info [<i>rp-address</i>] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about unacknowledged graft messages.	display pim ipv6 grafts [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the IPv6 PIM information on an interface or all interfaces.	display pim ipv6 interface [<i>interface-type interface-number</i>] [verbose] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about join/prune messages to send.	display pim ipv6 join-prune mode { sm [flags <i>flag-value</i>] ssm } [interface <i>interface-type interface-number</i> neighbor <i>ipv6-neighbor-address</i>] * [verbose] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 PIM neighboring information.	display pim ipv6 neighbor [interface <i>interface-type interface-number</i> <i>ipv6-neighbor-address</i> verbose] * [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about the IPv6 PIM routing table.	display pim ipv6 routing-table [<i>ipv6-group-address</i> [<i>prefix-length</i>] <i>ipv6-source-address</i> [<i>prefix-length</i>] incoming-interface [<i>interface-type interface-number</i> register] outgoing-interface { include exclude match } { <i>interface-type interface-number</i> register } mode <i>mode-type</i> flags <i>flag-value</i> fsm] * [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the RP information.	display pim ipv6 rp-info [<i>ipv6-group-address</i>] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Reset IPv6 PIM control message counters.	reset pim ipv6 control-message counters [interface <i>interface-type</i> <i>interface-number</i>]	Available in user view

IPv6 PIM configuration examples

IPv6 PIM-DM configuration example

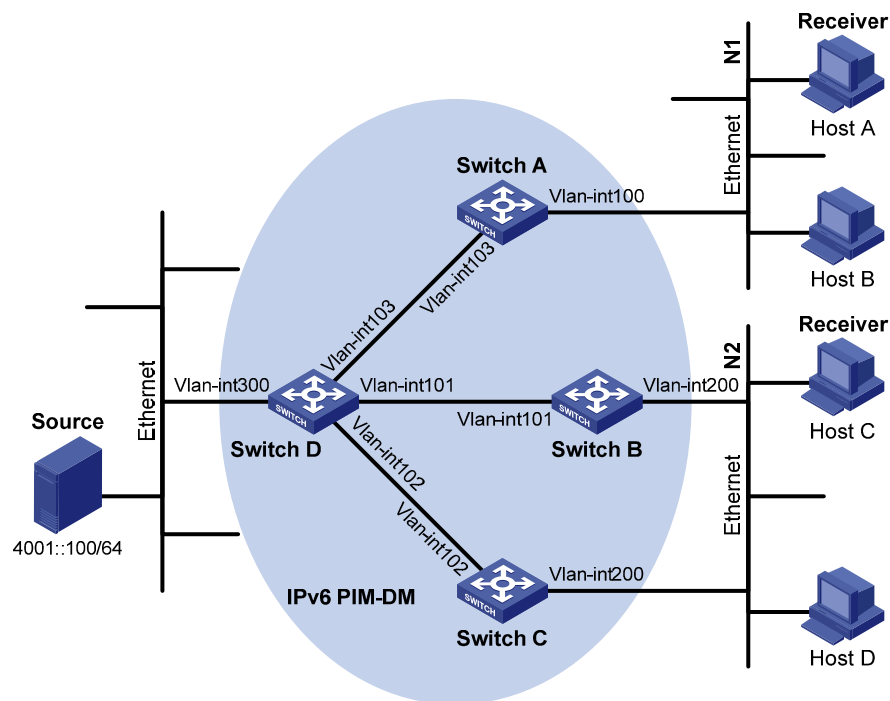
Network requirements

Receivers receive VOD information through multicast. The receiver groups of different organizations form stub networks, and one or more receiver hosts exist in each stub network. The entire IPv6 PIM domain is operating in the dense mode.

Host A and Host C are multicast receivers in two stub networks N1 and N2.

MLDv1 runs between Switch A and N1 and between Switch B/Switch C and N2.

Figure 103 Network diagram



Device	Interface	IPv6 address	Device	Interface	IPv6 address
Switch A	Vlan-int100	1001::1/64	Switch D	Vlan-int300	4001::1/64
	Vlan-int103	1002::1/64		Vlan-int103	1002::2/64
Switch B	Vlan-int200	2001::1/64	Switch D	Vlan-int101	2002::2/64
	Vlan-int101	2002::1/64		Vlan-int102	3001::2/64
Switch C	Vlan-int200	2001::2/64			
	Vlan-int102	3001::1/64			

Configuration procedure

1. Enable IPv6 forwarding and configure IPv6 addresses and IPv6 unicast routing:
Enable IPv6 forwarding on each switch and configure the IPv6 address and prefix length for each interface as per [Figure 103](#). (Details not shown.)
Configure OSPFv3 on the switches in the IPv6 PIM-DM domain to ensure network-layer reachability among them. (Details not shown.)

2. Enable IPv6 multicast routing, and enable IPv6 PIM-DM and MLD:

Enable IPv6 multicast routing on Switch A, enable IPv6 PIM-DM on each interface, and enable MLD on VLAN-interface 100, which connects Switch A to N1.

```
<SwitchA> system-view
[SwitchA] multicast ipv6 routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] mld enable
[SwitchA-Vlan-interface100] pim ipv6 dm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 103
[SwitchA-Vlan-interface103] pim ipv6 dm
[SwitchA-Vlan-interface103] quit
```

The configuration on Switch B and Switch C is similar to that on Switch A.

Enable IPv6 multicast routing on Switch D, and enable IPv6 PIM-DM on each interface.

```
<SwitchD> system-view
[SwitchD] multicast ipv6 routing-enable
[SwitchD] interface vlan-interface 300
[SwitchD-Vlan-interface300] pim ipv6 dm
[SwitchD-Vlan-interface300] quit
[SwitchD] interface vlan-interface 103
[SwitchD-Vlan-interface103] pim ipv6 dm
[SwitchD-Vlan-interface103] quit
[SwitchD] interface vlan-interface 101
[SwitchD-Vlan-interface101] pim ipv6 dm
[SwitchD-Vlan-interface101] quit
[SwitchD] interface vlan-interface 102
[SwitchD-Vlan-interface102] pim ipv6 dm
[SwitchD-Vlan-interface102] quit
```

3. Verify the configuration:

Use the **display pim ipv6 interface** command to view the IPv6 PIM configuration and running status on each interface. For example:

Display IPv6 PIM information on Switch D.

```
[SwitchD] display pim ipv6 interface
```

Interface	NbrCnt	HelloInt	DR-Pri	DR-Address
Vlan300	0	30	1	4001::1 (local)
Vlan103	0	30	1	1002::2 (local)
Vlan101	1	30	1	2002::2 (local)


```
Vlan102          1          30          1          3001::2
                (local)
```

Use the **display pim ipv6 neighbor** command to view the IPv6 PIM neighboring relationships among the switches. For example:

Display IPv6 PIM neighboring relationships on Switch D.

```
[SwitchD] display pim ipv6 neighbor
Total Number of Neighbors = 3
```

Neighbor	Interface	Uptime	Expires	Dr-Priority
1002::1	Vlan103	00:04:00	00:01:29	1
2002::1	Vlan101	00:04:16	00:01:29	3
3001::1	Vlan102	00:03:54	00:01:17	5

Assume that Host A needs to receive the information addressed to IPv6 multicast group G (FF0E::101). After IPv6 multicast source S (4001::100/64) sends IPv6 multicast packets to the IPv6 multicast group G, an SPT is established through traffic flooding. Switches on the SPT path (Switch A and Switch D) have their (S, G) entries. Host A sends an MLD report to Switch A to join IPv6 multicast group G, and a (*, G) entry is generated on Switch A. You can use the **display pim IPv6 routing-table** command to view the IPv6 PIM routing table information on each switch. For example:

Display IPv6 PIM multicast routing table information on Switch A.

```
[SwitchA] display pim ipv6 routing-table
Total 1 (*, G) entry; 1 (S, G) entry
```

```
(*, FF0E::101)
  Protocol: pim-dm, Flag: WC
  UpTime: 00:01:24
  Upstream interface: NULL
    Upstream neighbor: NULL
    RPF prime neighbor: NULL
  Downstream interface(s) information:
  Total number of downstreams: 1
    1: Vlan-interface100
      Protocol: mld, UpTime: 00:01:20, Expires: never

(4001::100, FF0E::101)
  Protocol: pim-dm, Flag: ACT
  UpTime: 00:01:20
  Upstream interface: Vlan-interface103
    Upstream neighbor: 1002::2
    RPF prime neighbor: 1002::2
  Downstream interface(s) information:
  Total number of downstreams: 1
    1: Vlan-interface100
      Protocol: pim-dm, UpTime: 00:01:20, Expires: never
```

Display IPv6 PIM multicast routing table information on Switch D.

```
[SwitchD] display pim ipv6 routing-table
Total 0 (*, G) entry; 1 (S, G) entry
```

```
(4001::100, FFOE::101)
  Protocol: pim-dm, Flag: LOC ACT
  UpTime: 00:02:19
  Upstream interface: Vlan-interface300
    Upstream neighbor: NULL
    RPF prime neighbor: NULL
  Downstream interface(s) information:
  Total number of downstreams: 2
    1: Vlan-interface103
      Protocol: pim-dm, UpTime: 00:02:19, Expires: never
    2: Vlan-interface102
      Protocol: pim-dm, UpTime: 00:02:19, Expires: never
```

IPv6 PIM-SM non-scoped zone configuration example

Network requirements

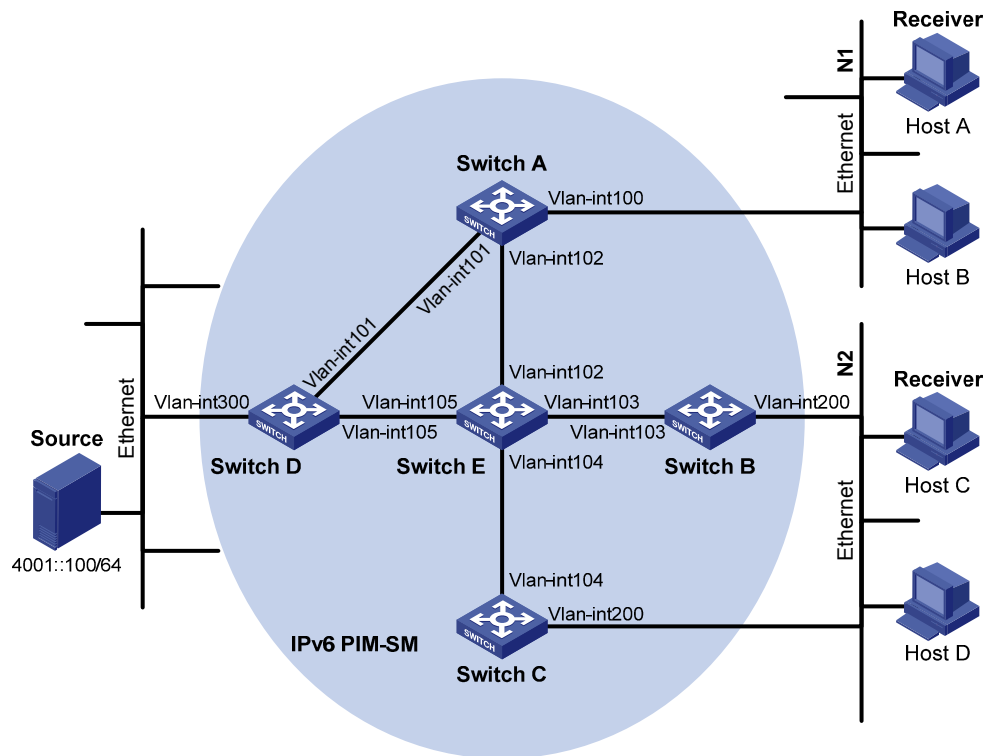
Receivers receive VOD information through multicast. The receiver groups of different organizations form stub networks, and one or more receiver hosts exist in each stub network. The entire PIM domain is operating in the sparse mode.

Host A and Host C are IPv6 multicast receivers in two stub networks, N1 and N2.

VLAN-interface 105 on Switch D and VLAN-interface 102 on Switch E act as C-BSRs and C-RPs. The C-BSR on Switch E has a higher priority. The IPv6 multicast group range served by the C-RP is FFOE::101/64. Modify the hash mask length to map a certain number of consecutive IPv6 group addresses within the range to the two C-RPs.

MLDv1 runs between Switch A and N1 and between Switch B/Switch C and N2.

Figure 104 Network diagram



Device	Interface	IPv6 address	Device	Interface	IPv6 address
Switch A	Vlan-int100	1001::1/64	Switch D	Vlan-int300	4001::1/64
	Vlan-int101	1002::1/64		Vlan-int101	1002::2/64
	Vlan-int102	1003::1/64		Vlan-int105	4002::1/64
Switch B	Vlan-int200	2001::1/64	Switch E	Vlan-int104	3001::2/64
	Vlan-int103	2002::1/64		Vlan-int103	2002::2/64
Switch C	Vlan-int200	2001::2/64		Vlan-int102	1003::2/64
	Vlan-int104	3001::1/64		Vlan-int105	4002::2/64

Configuration procedure

1. Enable IPv6 forwarding and configure IPv6 addresses and IPv6 unicast routing:
Enable IPv6 forwarding on each switch and configure the IPv6 address and prefix length for each interface as per Figure 104. (Details not shown.)

Configure OSPFv3 on the switches in the IPv6 PIM-DM domain to ensure network-layer reachability among them. (Details not shown.)

2. Enable IPv6 multicast routing, and enable IPv6 PIM-SM and MLD:

Enable IPv6 multicast routing on Switch A, enable IPv6 PIM-SM on each interface, and enable MLD on VLAN-interface 300, which connects Switch A to N1.

```
<SwitchA> system-view
[SwitchA] multicast ipv6 routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] mld enable
[SwitchA-Vlan-interface100] pim ipv6 sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
```

```
[SwitchA-Vlan-interface101] pim ipv6 sm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 102
[SwitchA-Vlan-interface102] pim ipv6 sm
[SwitchA-Vlan-interface102] quit
```

The configuration on Switch B and Switch C is similar to that on Switch A. The configuration on Switch D and Switch E is also similar to that on Switch A except that it is not necessary to enable MLD on the corresponding interfaces on these two switches.

3. Configure a C-BSR and a C-RP:

On Switch D, configure the service scope of RP advertisements, specify a C-BSR and a C-RP, and set the hash mask length to 128 and the priority of the C-BSR to 10.

```
<SwitchD> system-view
[SwitchD] acl ipv6 number 2005
[SwitchD-acl6-basic-2005] rule permit source ff0e::101 64
[SwitchD-acl6-basic-2005] quit
[SwitchD] pim ipv6
[SwitchD-pim6] c-bsr 4002::1 128 10
[SwitchD-pim6] c-rp 4002::1 group-policy 2005
[SwitchD-pim6] quit
```

On Switch E, configure the service scope of RP advertisements, specify a C-BSR and a C-RP, and set the hash mask length to 128 and the priority of the C-BSR to 20.

```
<SwitchE> system-view
[SwitchE] acl ipv6 number 2005
[SwitchE-acl6-basic-2005] rule permit source ff0e::101 64
[SwitchE-acl6-basic-2005] quit
[SwitchE] pim ipv6
[SwitchE-pim6] c-bsr 1003::2 128 20
[SwitchE-pim6] c-rp 1003::2 group-policy 2005
[SwitchE-pim6] quit
```

4. Verify the configuration:

Use the **display pim ipv6 interface** command to view the IPv6 PIM configuration and running status on each interface. For example:

Display IPv6 PIM information on all interfaces of Switch A.

```
[SwitchA] display pim ipv6 interface
Interface           NbrCnt  HelloInt  DR-Pri  DR-Address
Vlan100             0        30        1       1001::1
                   (local)
Vlan101             1        30        1       1002::2
Vlan102             1        30        1       1003::2
```

To view the BSR election information and the locally configured C-RP information in effect on a switch, use the **display pim ipv6 bsr-info** command. For example:

Display BSR information and the locally configured C-RP information in effect on Switch A.

```
[SwitchA] display pim ipv6 bsr-info
Elected BSR Address: 1003::2
Priority: 20
Hash mask length: 128
State: Accept Preferred
```

```
Uptime: 00:04:22
Expires: 00:01:46
```

Display BSR information and the locally configured C-RP information in effect on Switch D.

```
[SwitchD] display pim ipv6 bsr-info
```

```
Elected BSR Address: 1003::2
  Priority: 20
  Hash mask length: 128
  State: Elected
  Uptime: 00:05:26
  Expires: 00:01:45
Candidate BSR Address: 4002::1
  Priority: 10
  Hash mask length: 128
  State: Candidate
```

```
Candidate RP: 4002::1(Vlan-interface105)
  Priority: 192
  HoldTime: 130
  Advertisement Interval: 60
  Next advertisement scheduled at: 00:00:48
```

Display BSR information and the locally configured C-RP information in effect on Switch E.

```
[SwitchE] display pim ipv6 bsr-info
```

```
Elected BSR Address: 1003::2
  Priority: 20
  Hash mask length: 128
  State: Elected
  Uptime: 00:01:10
  Next BSR message scheduled at: 00:01:48
Candidate BSR Address: 1003::2
  Priority: 20
  Hash mask length: 128
  State: Elected
```

```
Candidate RP: 1003::2(Vlan-interface102)
  Priority: 192
  HoldTime: 130
  Advertisement Interval: 60
  Next advertisement scheduled at: 00:00:48
```

To view the RP information discovered on a switch, use the **display pim ipv6 rp-info** command. For example:

Display RP information on Switch A.

```
[SwitchA] display pim ipv6 rp-info
```

```
PIM-SM BSR RP information:
prefix/prefix length: FF0E::101/64
  RP: 4002::1
  Priority: 192
  HoldTime: 130
  Uptime: 00:05:19
```

```
Expires: 00:02:11
```

```
RP: 1003::2
```

```
Priority: 192
```

```
HoldTime: 130
```

```
Uptime: 00:05:19
```

```
Expires: 00:02:11
```

Assume that Host A needs to receive information addressed to the IPv6 multicast group G (FF0E::100). The RP corresponding to the multicast group G is Switch E as a result of hash calculation, so an RPT will be built between Switch A and Switch E. When the IPv6 multicast source S (4001::100/64) registers with the RP, an SPT will be built between Switch D and Switch E. After receiving IPv6 multicast data, Switch A immediately switches from the RPT to the SPT. The switches on the RPT path (Switch A and Switch E) have a (*, G) entry, and the switches on the SPT path (Switch A and Switch D) have an (S, G) entry. You can use the **display pim ipv6 routing-table** command to view the PIM routing table information on the switches. For example:

```
# Display IPv6 PIM multicast routing table information on Switch A.
```

```
[SwitchA] display pim ipv6 routing-table
```

```
Total 1 (*, G) entry; 1 (S, G) entry
```

```
(*, FF0E::100)
```

```
RP: 1003::2
```

```
Protocol: pim-sm, Flag: WC
```

```
UpTime: 00:03:45
```

```
Upstream interface: Vlan-interface102
```

```
Upstream neighbor: 1003::2
```

```
RPF prime neighbor: 1003::2
```

```
Downstream interface(s) information:
```

```
Total number of downstreams: 1
```

```
1: Vlan-interface100
```

```
Protocol: mld, UpTime: 00:02:15, Expires: 00:03:06
```

```
(4001::100, FF0E::100)
```

```
RP: 1003::2
```

```
Protocol: pim-sm, Flag: SPT ACT
```

```
UpTime: 00:02:15
```

```
Upstream interface: Vlan-interface101
```

```
Upstream neighbor: 1002::2
```

```
RPF prime neighbor: 1002::2
```

```
Downstream interface(s) information:
```

```
Total number of downstreams: 1
```

```
1: Vlan-interface100
```

```
Protocol: pim-sm, UpTime: 00:02:15, Expires: 00:03:06
```

```
# Display IPv6 PIM multicast routing table information on Switch D.
```

```
[SwitchD] display pim ipv6 routing-table
```

```
Total 0 (*, G) entry; 1 (S, G) entry
```

```
(4001::100, FF0E::100)
```

```
RP: 1003::2
```

```

Protocol: pim-sm, Flag: SPT LOC ACT
UpTime: 00:14:44
Upstream interface: Vlan-interface300
    Upstream neighbor: NULL
    RPF prime neighbor: NULL
Downstream interface(s) information:
Total number of downstreams: 1
    1: Vlan-interface105
        Protocol: mld, UpTime: 00:14:44, Expires: 00:02:26
# Display IPv6 PIM multicast routing table information on Switch E.
[SwitchE] display pim ipv6 routing-table
Total 1 (*, G) entry; 0 (S, G) entry

(*, FF0E::100)
RP: 1003::2 (local)
Protocol: pim-sm, Flag: WC
UpTime: 00:16:56
Upstream interface: Register
    Upstream neighbor: 4002::1
    RPF prime neighbor: 4002::1
Downstream interface(s) information:
Total number of downstreams: 1
    1: Vlan-interface102
        Protocol: pim-sm, UpTime: 00:16:56, Expires: 00:02:34

```

IPv6 PIM-SM admin-scope zone configuration example

Network requirements

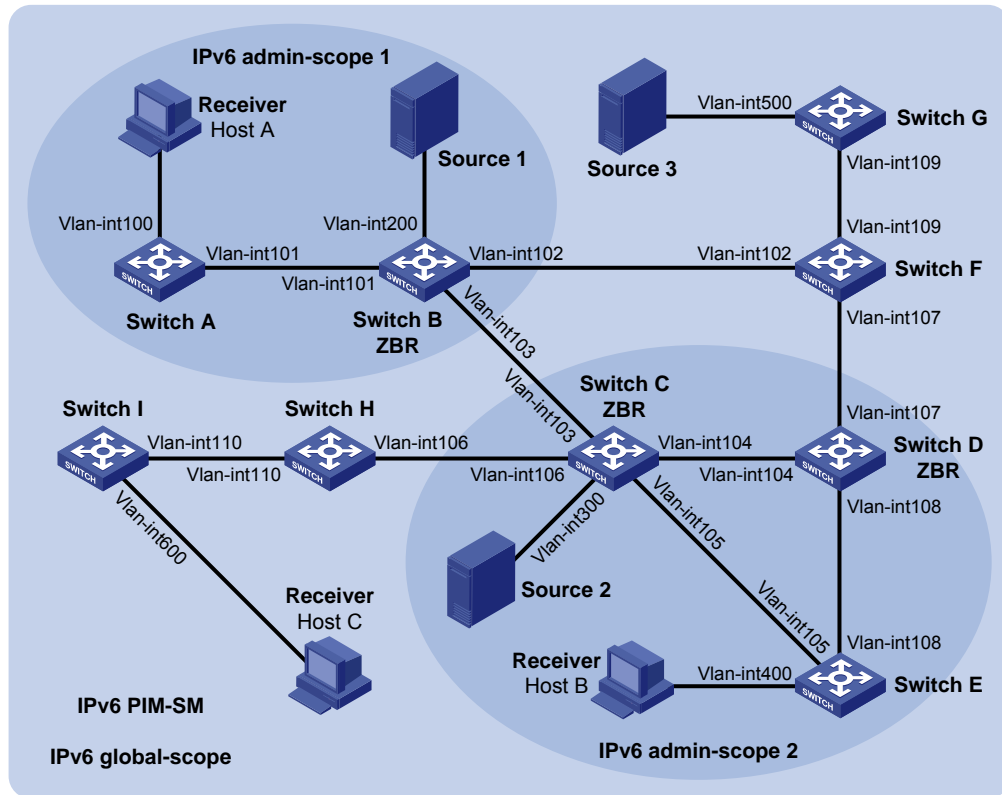
Receivers receive VOD information through multicast. The entire IPv6 PIM-SM domain is divided into IPv6 admin-scope zone 1, IPv6 admin-scope zone 2, and the IPv6 global zone. Switch B, Switch C, and Switch D are ZBRs of these three domains respectively.

Source 1 and Source 2 send different multicast information to FF14::101. Host A receives the multicast information only from Source 1, and Host B receives the multicast information only from Source 2. Source 3 sends multicast information to multicast group FF1E::202. Host C is a multicast receiver for this multicast group.

VLAN-interface 101 of Switch B acts as a C-BSR and C-RP of admin-scope zone 1, which serves the IPv6 multicast groups with the Scope field value in their group addresses being 4. VLAN-interface 104 of Switch D acts as a C-BSR and C-RP of admin-scope zone 2, which also serves the IPv6 multicast groups with the Scope field value in their group addresses being 4. VLAN-interface 109 of Switch F acts as C-BSRs and C-RPs of the global scope zone, which serves IPv6 multicast groups with the Scope field value in their group addresses being 14.

MLDv1 runs between Switch A, Switch E, Switch I, and their respective receivers.

Figure 105 Network diagram



Device	Interface	IPv6 address	Device	Interface	IPv6 address
Switch A	Vlan-int100	1001::1/64	Switch D	Vlan-int104	3002::2/64
	Vlan-int101	1002::1/64		Vlan-int108	6001::1/64
Switch B	Vlan-int200	2001::1/64		Vlan-int107	6002::1/64
	Vlan-int101	1002::2/64	Switch E	Vlan-int400	7001::1/64
	Vlan-int103	2002::1/64		Vlan-int105	3003::2/64
	Vlan-int102	2003::1/64		Vlan-int108	6001::2/64
Switch C	Vlan-int300	3001::1/64	Switch F	Vlan-int109	8001::1/64
	Vlan-int104	3002::1/64		Vlan-int107	6002::2/64
	Vlan-int105	3003::1/64		Vlan-int102	2003::2/64
	Vlan-int103	2002::2/64	Switch G	Vlan-int500	9001::1/64
	Vlan-int106	3004::1/64		Vlan-int109	8001::2/64
Switch H	Vlan-int110	4001::1/64	Source 1	—	2001::100/64
	Vlan-int106	3004::2/64	Source 2	—	3001::100/64
Switch I	Vlan-int600	5001::1/64	Source 3	—	9001::100/64
	Vlan-int110	4001::2/64			

Configuration procedure

1. Configure IPv6 addresses and unicast routing:

Configure the IPv6 address and prefix length for each interface as per Figure 105. (Details not shown.)

Configure OSPFv3 on the switches in the IPv6 PIM-SM domain to ensure network-layer reachability among them. (Details not shown.)

2. Enable IPv6 multicast routing and IPv6 administrative scoping, and enable IPv6 PIM-SM and MLD:
Enable IPv6 multicast routing and administrative scoping on Switch A, enable IPv6 PIM-SM on each interface, and enable MLD on the host-side interface VLAN-interface 100.

```
<SwitchA> system-view
[SwitchA] multicast ipv6 routing-enable
[SwitchA] pim ipv6
[SwitchA-pim6] c-bsr admin-scope
[SwitchA-pim6] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] mld enable
[SwitchA-Vlan-interface100] pim ipv6 sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim ipv6 sm
[SwitchA-Vlan-interface101] quit
```

The configuration on Switch E and Switch I is similar to the configuration on Switch A.

- # On Switch B, enable IPv6 multicast routing and IPv6 administrative scoping, and enable IPv6 PIM-SM on each interface.

```
<SwitchB> system-view
[SwitchB] multicast ipv6 routing-enable
[SwitchB] pim ipv6
[SwitchB-pim6] c-bsr admin-scope
[SwitchB-pim6] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] pim ipv6 sm
[SwitchB-Vlan-interface200] quit
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] pim ipv6 sm
[SwitchB-Vlan-interface101] quit
[SwitchB] interface vlan-interface 102
[SwitchB-Vlan-interface102] pim ipv6 sm
[SwitchB-Vlan-interface102] quit
[SwitchB] interface vlan-interface 103
[SwitchB-Vlan-interface103] pim ipv6 sm
[SwitchB-Vlan-interface103] quit
```

The configuration on Switch C, Switch D, Switch F, Switch G, and Switch H is similar to the configuration on Switch B. (Details not shown.)

3. Configure an admin-scope zone boundary:

- # On Switch B, configure VLAN-interface 102 and VLAN-interface 103 to be the boundary of admin-scope zone 1.

```
[SwitchB] interface vlan-interface 102
[SwitchB-Vlan-interface102] multicast ipv6 boundary scope 4
[SwitchB-Vlan-interface102] quit
[SwitchB] interface vlan-interface 103
[SwitchB-Vlan-interface103] multicast ipv6 boundary scope 4
[SwitchB-Vlan-interface103] quit
```

On Switch C, configure VLAN-interface 103 and VLAN-interface 106 to be the boundary of admin-scope zone 2.

```
<SwitchC> system-view
[SwitchC] interface vlan-interface 103
[SwitchC-Vlan-interface103] multicast ipv6 boundary scope 4
[SwitchC-Vlan-interface103] quit
[SwitchC] interface vlan-interface 106
[SwitchC-Vlan-interface106] multicast ipv6 boundary scope 4
[SwitchC-Vlan-interface106] quit
```

On Switch D, configure VLAN-interface 107 to be the boundary of admin-scope zone 2.

```
<SwitchD> system-view
[SwitchD] interface vlan-interface 107
[SwitchD-Vlan-interface107] multicast ipv6 boundary scope 4
[SwitchD-Vlan-interface107] quit
```

4. Configure C-BSRs and C-RPs:

On Switch B, configure the service scope of RP advertisements, and configure VLAN-interface 101 as a C-BSR and C-RP of admin-scope zone 1.

```
[SwitchB] pim ipv6
[SwitchB-pim6] c-bsr scope 4
[SwitchB-pim6] c-bsr 1002::2
[SwitchB-pim6] c-rp 1002::2 scope 4
[SwitchB-pim6] quit
```

On Switch D, configure the service scope of RP advertisements, and configure VLAN-interface 104 as a C-BSR and C-RP of admin-scope zone 2.

```
[SwitchD] pim ipv6
[SwitchD-pim6] c-bsr scope 4
[SwitchD-pim6] c-bsr 3002::2
[SwitchD-pim6] c-rp 3002::2 scope 4
[SwitchD-pim6] quit
```

On Switch F, configure VLAN-interface 109 as a C-BSR and C-RP in the global scope zone.

```
<SwitchF> system-view
[SwitchF] pim ipv6
[SwitchF-pim6] c-bsr scope global
[SwitchF-pim6] c-bsr 8001::1
[SwitchF-pim6] c-rp 8001::1
[SwitchF-pim6] quit
```

5. Verify the configuration:

To view the BSR election information and the C-RP information on a switch, use the **display pim ipv6 bsr-info** command. For example:

Display BSR information and the locally configured C-RP information on Switch B.

```
[SwitchB] display pim ipv6 bsr-info
Elected BSR Address: 8001::1
    Priority: 64
    Hash mask length: 126
    State: Accept Preferred
    Scope: 14
    Uptime: 00:01:45
```

```
Expires: 00:01:25
Elected BSR Address: 1002::2
  Priority: 64
  Hash mask length: 126
  State: Elected
  Scope: 4
  Uptime: 00:04:54
  Next BSR message scheduled at: 00:00:06
```

```
Candidate BSR Address: 1002::2
  Priority: 64
  Hash mask length: 126
  State: Elected
  Scope: 4
```

```
Candidate RP: 1002::2(Vlan-interface101)
  Priority: 192
  HoldTime: 130
  Advertisement Interval: 60
  Next advertisement scheduled at: 00:00:15
```

Display BSR information and the locally configured C-RP information on Switch D.

```
[SwitchD] display pim ipv6 bsr-info
```

```
Elected BSR Address: 8001::1
  Priority: 64
  Hash mask length: 126
  State: Accept Preferred
  Scope: 14
  Uptime: 00:01:45
  Expires: 00:01:25
```

```
Elected BSR Address: 3002::2
  Priority: 64
  Hash mask length: 126
  State: Elected
  Scope: 4
  Uptime: 00:03:48
  Next BSR message scheduled at: 00:01:12
```

```
Candidate BSR Address: 3002::2
  Priority: 64
  Hash mask length: 126
  State: Elected
  Scope: 4
```

```
Candidate RP: 3002::2(Vlan-interface104)
  Priority: 192
  HoldTime: 130
  Advertisement Interval: 60
  Next advertisement scheduled at: 00:00:10
```

Display BSR information and the locally configured C-RP information on Switch F.

```
[SwitchF] display pim ipv6 bsr-info
```

```
Elected BSR Address: 8001::1
  Priority: 64
  Hash mask length: 126
  State: Elected
  Scope: 14
  Uptime: 00:01:11
  Next BSR message scheduled at: 00:00:49
```

```
Candidate BSR Address: 8001::1
  Priority: 64
  Hash mask length: 126
  State: Elected
  Scope: 14
```

```
Candidate RP: 8001::1(Vlan-interface109)
  Priority: 192
  HoldTime: 130
  Advertisement Interval: 60
  Next advertisement scheduled at: 00:00:55
```

To view the RP information learned on a switch, use the **display pim ipv6 rp-info** command. For example:

```
# Display RP information on Switch B.
```

```
[SwitchB] display pim ipv6 rp-info
```

```
PIM-SM BSR RP information:
```

```
prefix/prefix length: FF0E::/16
  RP: 8001::1
  Priority: 192
  HoldTime: 130
  Uptime: 00:03:39
  Expires: 00:01:51
```

```
prefix/prefix length: FF1E::/16
  RP: 8001::1
  Priority: 192
  HoldTime: 130
  Uptime: 00:03:39
  Expires: 00:01:51
```

```
prefix/prefix length: FF2E::/16
  RP: 8001::1
  Priority: 192
  HoldTime: 130
  Uptime: 00:03:39
  Expires: 00:01:51
```

```
prefix/prefix length: FF3E::/16
  RP: 8001::1
  Priority: 192
  HoldTime: 130
```

Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF4E::/16
RP: 8001::1
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF5E::/16
RP: 8001::1
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF6E::/16
RP: 8001::1
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF7E::/16
RP: 8001::1
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF8E::/16
RP: 8001::1
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF9E::/16
RP: 8001::1
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FFAE::/16
RP: 8001::1
Priority: 192

HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FFBE::/16
RP: 8001::1
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FFCE::/16
RP: 8001::1
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FFDE::/16
RP: 8001::1
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FFEE::/16
RP: 8001::1
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FFFE::/16
RP: 8001::1
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF04::/16
RP: 1002::2
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF14::/16
RP: 1002::2

Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF24::/16
RP: 1002::2
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF34::/16
RP: 1002::2
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF44::/16
RP: 1002::2
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF54::/16
RP: 1002::2
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF64::/16
RP: 1002::2
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF74::/16
RP: 1002::2
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF84::/16

RP: 1002::2
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF94::/16

RP: 1002::2
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FFA4::/16

RP: 1002::2
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FFB4::/16

RP: 1002::2
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FFC4::/16

RP: 1002::2
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FFD4::/16

RP: 1002::2
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FFE4::/16

RP: 1002::2
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51


```
prefix/prefix length: FFF4::/16
  RP: 1002::2
  Priority: 192
  HoldTime: 130
  Uptime: 00:03:39
  Expires: 00:01:51
```

Display RP information on Switch F.

```
[SwitchF] display pim rp-info
PIM-SM BSR RP information:
prefix/prefix length: FF0E::/16
  RP: 8001::1
  Priority: 192
  HoldTime: 130
  Uptime: 00:03:39
  Expires: 00:01:51
```

```
prefix/prefix length: FF1E::/16
  RP: 8001::1
  Priority: 192
  HoldTime: 130
  Uptime: 00:03:39
  Expires: 00:01:51
```

```
prefix/prefix length: FF2E::/16
  RP: 8001::1
  Priority: 192
  HoldTime: 130
  Uptime: 00:03:39
  Expires: 00:01:51
```

```
prefix/prefix length: FF3E::/16
  RP: 8001::1
  Priority: 192
  HoldTime: 130
  Uptime: 00:03:39
  Expires: 00:01:51
```

```
prefix/prefix length: FF4E::/16
  RP: 8001::1
  Priority: 192
  HoldTime: 130
  Uptime: 00:03:39
  Expires: 00:01:51
```

```
prefix/prefix length: FF5E::/16
  RP: 8001::1
  Priority: 192
  HoldTime: 130
```

Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF6E::/16
RP: 8001::1
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF7E::/16
RP: 8001::1
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF8E::/16
RP: 8001::1
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FF9E::/16
RP: 8001::1
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FFAE::/16
RP: 8001::1
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FFBE::/16
RP: 8001::1
Priority: 192
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51

prefix/prefix length: FFCE::/16
RP: 8001::1
Priority: 192

```
HoldTime: 130
Uptime: 00:03:39
Expires: 00:01:51
```

```
prefix/prefix length: FFDE::/16
  RP: 8001::1
  Priority: 192
  HoldTime: 130
  Uptime: 00:03:39
  Expires: 00:01:51
```

```
prefix/prefix length: FFEE::/16
  RP: 8001::1
  Priority: 192
  HoldTime: 130
  Uptime: 00:03:39
  Expires: 00:01:51
```

```
prefix/prefix length: FFFE::/16
  RP: 8001::1
  Priority: 192
  HoldTime: 130
  Uptime: 00:03:39
  Expires: 00:01:51
```

IPv6 BIDIR-PIM configuration example

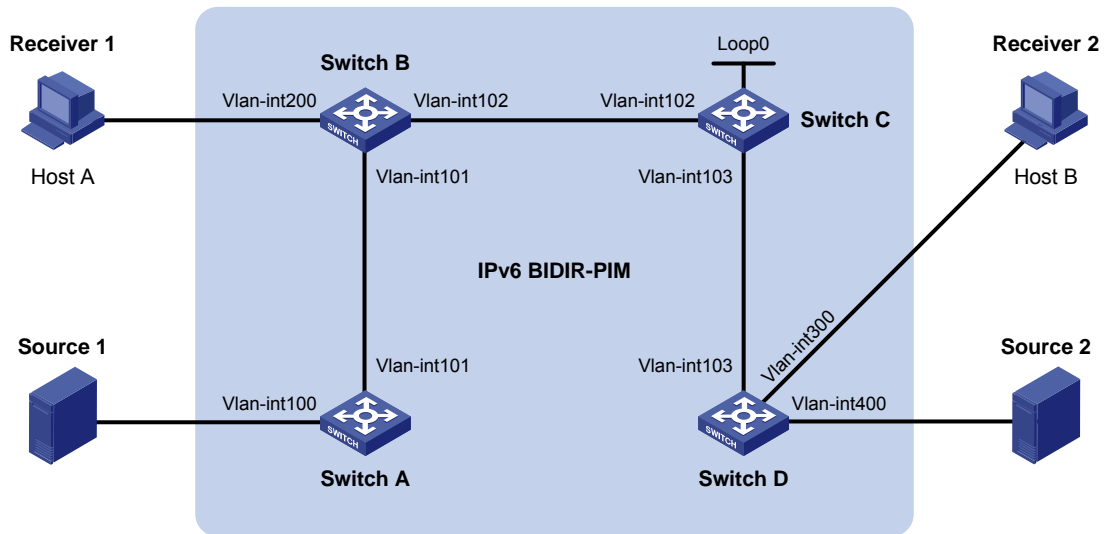
Network requirements

In the IPv6 BIDIR-PIM domain shown in [Figure 106](#). Source 1 and Source 2 send different IPv6 multicast information to IPv6 multicast group FF14::101. Host A and Host B receive IPv6 multicast information from the two sources.

VLAN interface 102 of Switch C acts as a C-BSR, and loopback interface 0 acts as a C-RP of the IPv6 BIDIR-PIM domain.

MLDv1 runs between Switch B and Host A and between Switch D and Host B.

Figure 106 Network diagram



Device	Interface	IPv6 address	Device	Interface	IPv6 address
Switch A	Vlan-int100	1001::1/64	Switch D	Vlan-int300	4001::1/64
	Vlan-int101	1002::1/64		Vlan-int400	5001::1/64
Switch B	Vlan-int200	2001::1/64	Source 1	-	1001::2/64
	Vlan-int101	1002::2/64		Source 2	-
Switch C	Vlan-int102	2002::1/64	Receiver 1	-	2001::2/64
	Vlan-int103	3001::1/64	Receiver 2	-	4001::2/64
	Loop0	6001::1/128			

Configuration procedure

1. Configure IPv6 forwarding, IPv6 addresses and the IPv6 unicast routing protocol:
 Enable IPv6 forwarding on each switch, and configure the IPv6 address and prefix length for each interface as per Figure 106. (Details not shown.)
 Configure OSPFv3 on the switches in the IPv6 BIDIR-PIM domain to ensure network-layer reachability among them. (Details not shown.)
2. Enable IPv6 multicast routing, IPv6 PIM-SM, IPv6 BIDIR-PIM, and MLD:
 # On Switch A, enable IPv6 multicast routing, enable IPv6 PIM-SM on each interface, and enable IPv6 BIDIR-PIM.

```

<SwitchA> system-view
[SwitchA] multicast ipv6 routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] pim ipv6 sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim ipv6 sm
[SwitchA-Vlan-interface101] quit
[SwitchA] pim ipv6
[SwitchA-pim6] bidir-pim enable
[SwitchA-pim6] quit
    
```

On Switch B, enable IPv6 multicast routing, enable IPv6 PIM-SM on each interface, enable MLD in VLAN interface 200, and enable IPv6 BIDIR-PIM.

```
<SwitchB> system-view
[SwitchB] multicast ipv6 routing-enable
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] mld enable
[SwitchB-Vlan-interface200] pim ipv6 sm
[SwitchB-Vlan-interface200] quit
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] pim ipv6 sm
[SwitchB-Vlan-interface101] quit
[SwitchB] interface vlan-interface 102
[SwitchB-Vlan-interface102] pim ipv6 sm
[SwitchB-Vlan-interface102] quit
[SwitchB] pim ipv6
[SwitchB-pim6] bidir-pim enable
[SwitchB-pim6] quit
```

On Switch C, enable IPv6 multicast routing, enable IPv6 PIM-SM on each interface, and enable IPv6 BIDIR-PIM.

```
<SwitchC> system-view
[SwitchC] multicast ipv6 routing-enable
[SwitchC] interface vlan-interface 102
[SwitchC-Vlan-interface102] pim ipv6 sm
[SwitchC-Vlan-interface102] quit
[SwitchC] interface vlan-interface 103
[SwitchC-Vlan-interface103] pim ipv6 sm
[SwitchC-Vlan-interface103] quit
[SwitchC] interface loopback 0
[SwitchC-LoopBack0] pim ipv6 sm
[SwitchC-LoopBack0] quit
[SwitchC] pim ipv6
[SwitchC-pim6] bidir-pim enable
```

On Switch D, enable IPv6 multicast routing, enable IPv6 PIM-SM on each interface, enable MLD in VLAN interface 300, and enable IPv6 BIDIR-PIM.

```
<SwitchD> system-view
[SwitchD] multicast ipv6 routing-enable
[SwitchD] interface vlan-interface 300
[SwitchD-Vlan-interface300] mld enable
[SwitchD-Vlan-interface300] pim ipv6 sm
[SwitchD-Vlan-interface300] quit
[SwitchD] interface vlan-interface 400
[SwitchD-Vlan-interface400] pim ipv6 sm
[SwitchD-Vlan-interface400] quit
[SwitchD] interface vlan-interface 103
[SwitchD-Vlan-interface103] pim ipv6 sm
[SwitchD-Vlan-interface103] quit
[SwitchD] pim ipv6
[SwitchD-pim6] bidir-pim enable
```

```
[SwitchD-pim6] quit
```

3. Configure C-BSR and C-RP:

On Switch C, configure VLAN interface 102 as a C-BSR, and loopback interface 0 as a C-RP for the entire IPv6 BIDIR-PIM domain.

```
[SwitchC-pim6] c-bsr 2002::2
```

```
[SwitchC-pim6] c-rp 6001::1 bidir
```

```
[SwitchC-pim6] quit
```

4. Verify the configuration:

To view the DF information of IPv6 BIDIR-PIM on a switch, use the **display pim ipv6 df-info** command:

Display the DF information of IPv6 BIDIR-PIM on Switch A.

```
[SwitchA] display pim ipv6 df-info
```

```
RP Address: 6001::1
```

Interface	State	DF-Pref	DF-Metric	DF-Uptime	DF-Address
Vlan100	Win	100	2	01:08:50	FE80::200:5EFF: FE71:2800 (local)
Vlan101	Lose	100	1	01:07:49	FE80::20F:E2FF: FE38:4E01

Display the DF information of IPv6 BIDIR-PIM on Switch B.

```
[SwitchB] display pim ipv6 df-info
```

```
RP Address: 6001::1
```

Interface	State	DF-Pref	DF-Metric	DF-Uptime	DF-Address
Vlan200	Win	100	1	01:24:09	FE80::200:5EFF: FE71:2801 (local)
Vlan101	Win	100	1	01:24:09	FE80::20F:E2FF: FE38:4E01 (local)
Vlan102	Lose	0	0	01:23:12	FE80::20F:E2FF: FE15:5601

Display the DF information of IPv6 BIDIR-PIM on Switch C.

```
[SwitchC] display pim ipv6 df-info
```

```
RP Address: 6001::1
```

Interface	State	DF-Pref	DF-Metric	DF-Uptime	DF-Address
Loop0	-	-	-	-	-
Vlan102	Win	0	0	01:06:07	FE80::20F:E2FF: FE15:5601 (local)
Vlan103	Win	0	0	01:06:07	FE80::20F:E2FF: FE15:5602 (local)

Display the DF information of IPv6 BIDIR-PIM on Switch D.

```
[SwitchD] display pim ipv6 df-info
```

```
RP Address: 6001::1
```

Interface	State	DF-Pref	DF-Metric	DF-Uptime	DF-Address
Vlan300	Win	100	1	01:19:53	FE80::200:5EFF: FE71:2803 (local)
Vlan400	Win	100	1	00:39:34	FE80::200:5EFF: FE71:2802 (local)
Vlan103	Lose	0	0	01:21:40	FE80::20F:E2FF: FE15:5602

To view the DF information of the IPv6 multicast forwarding table on a switch, use the **display multicast ipv6 forwarding-table df-info** command. For more information about this command, see *IP Multicast Command Reference*.

Display the DF information of the IPv6 multicast forwarding table on Switch A.

```
[SwitchA] display multicast ipv6 forwarding-table df-info
```

```
Multicast DF information
```

```
Total 1 RP
```

```
Total 1 RP matched
```

```
00001. RP Address: 6001::1
```

```
    MID: 0, Flags: 0x2100000:0
```

```
    Uptime: 00:08:32
```

```
    RPF interface: Vlan-interface101
```

```
    List of 1 DF interfaces:
```

```
        1: Vlan-interface100
```

Display the DF information of the IPv6 multicast forwarding table on Switch B.

```
[SwitchB] display multicast ipv6 forwarding-table df-info
```

```
Multicast DF information
```

```
Total 1 RP
```

```
Total 1 RP matched
```

```
00001. RP Address: 6001::1
```

```
    MID: 0, Flags: 0x2100000:0
```

```
    Uptime: 00:06:24
```

```
    RPF interface: Vlan-interface102
```

```
    List of 2 DF interfaces:
```

```
        1: Vlan-interface101
```

```
        2: Vlan-interface200
```

Display the DF information of the IPv6 multicast forwarding table on Switch C.

```
[SwitchC] display multicast ipv6 forwarding-table df-info
```

```
Multicast DF information
```

```
Total 1 RP
```

```
Total 1 RP matched
```

```
00001. RP Address: 6001::1
```

```
    MID: 0, Flags: 0x2100000:0
```

```
    Uptime: 00:07:21
```

```
    RPF interface: LoopBack0
```

```
    List of 2 DF interfaces:
```

```
        1: Vlan-interface102
```

```
        2: Vlan-interface103
```

Display the DF information of the IPv6 multicast forwarding table on Switch D.

```
[SwitchD] display multicast ipv6 forwarding-table df-info
```

```
Multicast DF information
```

```
Total 1 RP
```

```

Total 1 RP matched

00001. RP Address: 6001::1
MID: 0, Flags: 0x2100000:0
Uptime: 00:05:12
RPF interface: Vlan-interface103
List of 2 DF interfaces:
  1: Vlan-interface300
  2: Vlan-interface400

```

IPv6 PIM-SSM configuration example

Network requirements

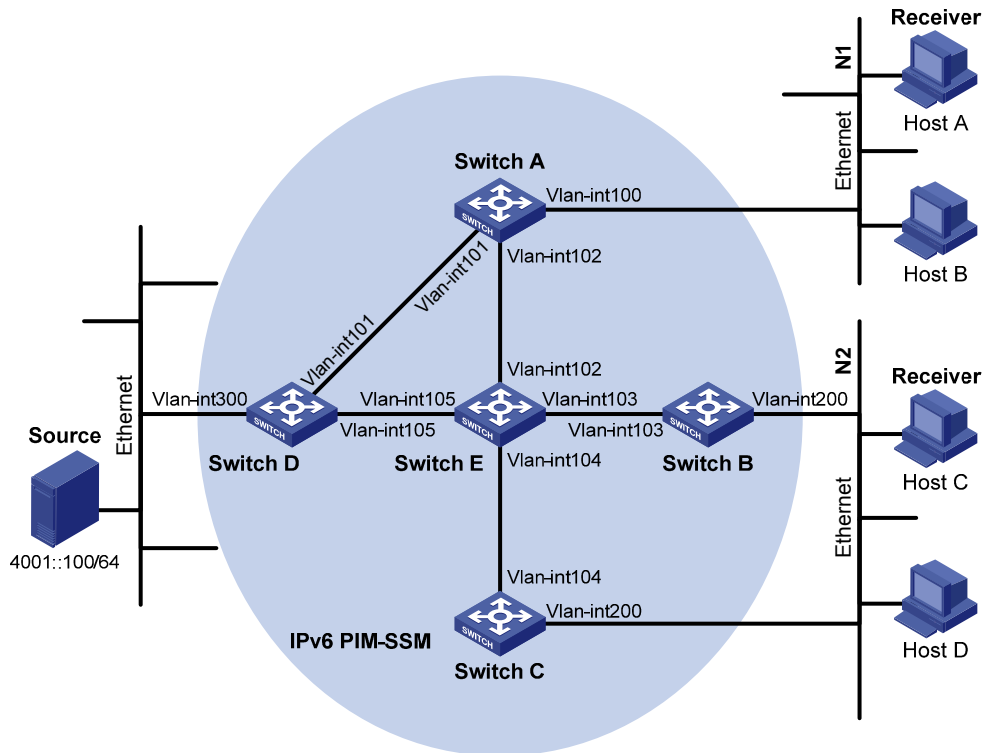
Receivers receive VOD information through multicast. The receiver groups of different organizations form stub networks, and one or more receiver hosts exist in each stub network. The entire PIM domain is operating in the SSM mode.

Host A and Host C are IPv6 multicast receivers in two stub networks, N1 and N2.

The SSM group range is FF3E::/64.

MLDv2 runs between Switch A and N1 and between Switch B/Switch C and N2.

Figure 107 Network diagram



Device	Interface	IPv6 address	Device	Interface	IPv6 address
Switch A	Vlan-int100	1001::1/64	Switch D	Vlan-int300	4001::1/64
	Vlan-int101	1002::1/64		Vlan-int101	1002::2/64
	Vlan-int102	1003::1/64		Vlan-int105	4002::1/64

Switch B	Vlan-int200	2001::1/64	Switch E	Vlan-int104	3001::2/64
	Vlan-int103	2002::1/64		Vlan-int103	2002::2/64
Switch C	Vlan-int200	2001::2/64		Vlan-int102	1003::2/64
	Vlan-int104	3001::1/64		Vlan-int105	4002::2/64

Configuration procedure

1. Enable IPv6 forwarding and configure IPv6 addresses and IPv6 unicast routing:
Enable IPv6 forwarding on each switch and configure the IPv6 address and prefix length for each interface as per [Figure 107](#). (Details not shown.)

Configure OSPFv3 on the switches in the IPv6 PIM-SSM domain to ensure network-layer reachability among them. (Details not shown.)

2. Enable IPv6 multicast routing, and enable IPv6 PIM-SM and MLD:

Enable IPv6 multicast routing on Switch A, enable IPv6 PIM-SM on each interface, and run MLDv2 on VLAN-interface 100, which connects Switch A to N1.

```
<SwitchA> system-view
[SwitchA] multicast ipv6 routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] mld enable
[SwitchA-Vlan-interface100] mld version 2
[SwitchA-Vlan-interface100] pim ipv6 sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim ipv6 sm
[SwitchA-Vlan-interface101] quit
[SwitchA] interface vlan-interface 102
[SwitchA-Vlan-interface102] pim ipv6 sm
[SwitchA-Vlan-interface102] quit
```

The configuration on Switch B and Switch C is similar to that on Switch A. The configuration on Switch D and Switch E is also similar to that on Switch A except that it is not necessary to enable MLD on the corresponding interfaces on these two switches.

3. Configure the IPv6 SSM group range:

Configure the IPv6 SSM group range to be FF3E::/64 on Switch A.

```
[SwitchA] acl ipv6 number 2000
[SwitchA-acl6-basic-2000] rule permit source ff3e:: 64
[SwitchA-acl6-basic-2000] quit
[SwitchA] pim ipv6
[SwitchA-pim6] ssm-policy 2000
[SwitchA-pim6] quit
```

The configuration on Switch B, Switch C, Switch D, and Switch E is similar to that on Switch A.

4. Verify the configuration:

Use the **display pim ipv6 interface** command to view the IPv6 PIM configuration and running status on each interface. For example:

Display IPv6 PIM configuration information on Switch A.

```
[SwitchA] display pim ipv6 interface
Interface          NbrCnt HelloInt   DR-Pri   DR-Address
Vlan100            0         30        1        1001::1
```

				(local)
Vlan101	1	30	1	1002::2
Vlan102	1	30	1	1003::2

Assume that Host A needs to receive the information a specific IPv6 multicast source S (4001::100/64) sends to IPv6 multicast group G (FF3E::101). Switch A builds an SPT toward the IPv6 multicast source. Switches on the SPT path (Switch A and Switch D) have generated an (S, G) entry, but Switch E, which is not on the SPT path, does not have IPv6 multicast routing entries. You can use the **display pim ipv6 routing-table** command to view the IPv6 PIM routing table information on each switch. For example:

Display IPv6 PIM multicast routing table information on Switch A.

```
[SwitchA] display pim ipv6 routing-table
Total 0 (*, G) entry; 1 (S, G) entry

(4001::100, FF3E::101)
  Protocol: pim-ssm, Flag:
  UpTime: 00:00:11
  Upstream interface: Vlan-interface101
    Upstream neighbor: 1002::2
    RPF prime neighbor: 1002::2
  Downstream interface(s) information:
  Total number of downstreams: 1
    1: Vlan-interface100
      Protocol: mld, UpTime: 00:00:11, Expires: 00:03:25
```

Display IPv6 PIM multicast routing table information on Switch B.

```
[SwitchD] display pim ipv6 routing-table
Total 0 (*, G) entry; 1 (S, G) entry

(4001::100, FF3E::101)
  Protocol: pim-ssm, Flag: LOC
  UpTime: 00:08:02
  Upstream interface: Vlan-interface300
    Upstream neighbor: NULL
    RPF prime neighbor: NULL
  Downstream interface(s) information:
  Total number of downstreams: 1
    1: Vlan-interface105
      Protocol: pim-ssm, UpTime: 00:08:02, Expires: 00:03:25
```

Troubleshooting IPv6 PIM configuration

Failure to build a multicast distribution tree correctly

Symptom

None of the routers in the network (including routers directly connected with IPv6 multicast sources and receivers) have IPv6 multicast forwarding entries. That is, a multicast distribution tree cannot be built correctly and clients cannot receive IPv6 multicast data.

Analysis

- An IPv6 PIM routing entry is created based on an IPv6 unicast route, whichever IPv6 PIM mode is running. Multicast works only when unicast does.
- IPv6 PIM must be enabled on the RPF interface. An RPF neighbor must be an IPv6 PIM neighbor as well. If IPv6 PIM is not enabled on the RPF interface or the RPF neighbor, the establishment of a multicast distribution tree will fail, resulting in abnormal multicast forwarding.
- IPv6 PIM requires that the same IPv6 PIM mode (namely, DM or SM) must run on the entire network. Otherwise, the establishment of a multicast distribution tree will fail, resulting in abnormal multicast forwarding.

Solution

1. Use the **display ipv6 routing-table** command to verify that a unicast route exists to the IPv6 multicast source or the RP.
2. Use the **display pim ipv6 interface** command to verify that the RPF interface is IPv6 PIM enabled. If IPv6 PIM is not enabled on the interface, use the **pim ipv6 dm** or **pim ipv6 sm** command to enable IPv6 PIM.
3. Use the **display pim ipv6 neighbor** command to verify that the RPF neighbor is an IPv6 PIM neighbor.
4. Verify that IPv6 PIM and MLD are enabled on the interfaces that are directly connected to the IPv6 multicast source and to the receiver.
5. Use the **display pim ipv6 interface verbose** command to verify that the same PIM mode is enabled on the RPF interface and the corresponding interface of the RPF neighbor router.
6. Use the **display current-configuration** command to verify that the same IPv6 PIM mode is enabled on all the routers in the entire network. Make sure that the same IPv6 PIM mode is enabled on all the routers: IPv6 PIM-SM on all routers, or IPv6 PIM-DM on all routers.

IPv6 multicast data abnormally terminated on an intermediate router

Symptom

An intermediate router can receive IPv6 multicast data successfully, but the data cannot reach the last hop router. An interface on the intermediate router receives data, but no corresponding (S, G) entry is created in the IPv6 PIM routing table.

Analysis

- If an IPv6 multicast forwarding boundary has been configured through the **multicast ipv6 boundary** command, any IPv6 multicast packet will be kept from crossing the boundary, and no routing entry can be created in the IPv6 PIM routing table.
- In addition, the **source-policy** command filters received IPv6 multicast packets. If the IPv6 multicast data fails to pass the IPv6 ACL rule defined in this command, IPv6 PIM cannot create the route entry either.

Solution

1. Use the **display current-configuration** command to verify the IPv6 multicast forwarding boundary settings. Use the **multicast ipv6 boundary** command to change the IPv6 multicast forwarding boundary settings.

2. Use the **display current-configuration** command to verify the IPv6 multicast filter configuration. Change the IPv6 ACL rule defined in the **source-policy** command so that the source/group address of the IPv6 multicast data can pass ACL filtering.

RPS cannot join SPT in IPv6 PIM-SM

Symptom

An RPT cannot be established correctly, or the RPs cannot join the SPT to the IPv6 multicast source.

Analysis

- As the core of an IPv6 PIM-SM domain, the RPs serves specific IPv6 multicast groups. Multiple RPs can coexist in a network. Make sure that the RP information on all routers is exactly the same and that a specific group is mapped to the same RP. Otherwise, IPv6 multicast will fail.
- In the case of the static RP mechanism, the same RP address must be configured on all the routers in the entire network, including static RPs, by means of the static RP command. Otherwise, IPv6 multicast will fail.

Solution

1. Use the **display ipv6 routing-table** command to verify that a route is available on each router to the RP.
2. Use the **display pim ipv6 rp-info** command to verify that the RP information is consistent on all routers. In the case of inconsistent RP information, configure consistent RP address on all the routers.
3. Use the **display pim ipv6 rp-info** command to verify that the same RP address has been configured on all the routers throughout the network.

RPT establishment failure or source registration failure in IPv6 PIM-SM

Symptom

C-RPs cannot unicast advertise messages to the BSR. The BSR does not advertise bootstrap messages containing C-RP information and has no unicast route to any C-RP. An RPT cannot be established correctly, or the DR cannot perform source registration with the RP.

Analysis

- C-RPs periodically send advertisement messages to the BSR by unicast. If a C-RP does not have a route to the BSR, the BSR cannot receive the advertisements from the C-RP, and the bootstrap messages of the BSR will not contain the information about that C-RP.
- The RP is the core of an IPv6 PIM-SM domain. Make sure that the RP information on all routers is exactly the same, a specific group is mapped to the same RP, and a unicast route is available to the RP.

Solution

1. Use the **display ipv6 routing-table** command to verify that routes are available on each router to the RP and the BSR, and whether a route is available between the RP and the BSR. Make sure that each C-RP has a unicast route to the BSR, the BSR has a unicast route to each C-RP, and all the routers in the entire network have a unicast route to the RP.

2. IPv6 PIM-SM needs the support of the RP and BSR. Use the **display pim ipv6 bsr-info** command to verify that the BSR information is available on each router, and then use the **display pim ipv6 rp-info** command to check whether the RP information is correct.
3. Use the **display pim ipv6 neighbor** command to verify that normal neighboring relationships have been established among the routers.

Configuring IPv6 MBGP (available only on the HP 5500 EI)

IPv6 MBGP overview

This chapter covers configuration tasks related to multiprotocol BGP for IPv6 multicast. For information about BGP and IPv6 BGP, see *Layer 3—IP Routing Configuration Guide*. The term "router" in this chapter refers to both routers and Layer 3 switches.

IETF defined Multiprotocol BGP (MP-BGP) to carry routing information for multiple network-layer protocols.

For an IPv6 network, the topology for IPv6 multicast might be different from that for IPv6 unicast. To distinguish them, the MP-BGP enables BGP to carry the IPv6 unicast Network Layer Reachability Information (NLRI) and IPv6 multicast NLRI separately, and the multicast NLRI performs reverse path forwarding (RPF) exclusively. In this way, route selection for a destination through the IPv6 unicast routing table and through the IPv6 multicast routing table have different results, ensuring consistent unicast forwarding and normal multicast between domains. For information about RPF, see "[Configuring multicast routing and forwarding \(available only on the HP 5500 EI\)](#)."

MP-BGP is defined in RFC 2858 (Multiprotocol Extensions for BGP-4). The application of MP-BGP on IPv6 multicast is called IPv6 Multicast BGP (IPv6 MBGP).

IPv6 MBGP configuration task list

Task	Remarks
Configuring basic IPv6 MBGP functions	Configuring an IPv6 MBGP peer Required
	Configuring a preferred value for routes from a peer or a peer group Optional
Controlling route distribution and reception	Injecting a local IPv6 MBGP route Optional
	Configuring IPv6 MBGP route redistribution Optional
	Configuring IPv6 MBGP route summarization Optional
	Advertising a default route to a peer or peer group Optional
	Configuring outbound IPv6 MBGP route filtering Optional
	Configuring inbound IPv6 MBGP route filtering Optional
	Configuring IPv6 MBGP route dampening Optional
Configuring IPv6 MBGP route attributes	Configuring IPv6 MBGP route preferences Optional
	Configuring the default local preference Optional
	Configuring the MED attribute Optional
	Configuring the next hop attribute Optional
Tuning and	Configuring the as_path attribute Optional
	Configuring IPv6 MBGP soft reset Optional

Task		Remarks
optimizing IPv6 MBGP networks	Enabling the IPv6 MBGP orf capability	Optional
	Configuring the maximum number of equal-cost routes for load-balancing	Optional
Configuring a large scale IPv6 MBGP network	Configuring an IPv6 MBGP peer group	Optional
	Configuring IPv6 MBGP community	Optional
	Configuring an IPv6 MBGP route reflector	Optional

Configuring basic IPv6 MBGP functions

Configuration prerequisites

IPv6 MBGP is an application of multiprotocol BGP. Before you configure IPv6 MBGP, complete the following tasks:

- Enable IPv6.
- Configure network layer addresses for interfaces.
- Complete BGP basic configuration.

Configuring an IPv6 MBGP peer

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable BGP and enter BGP view.	bgp <i>as-number</i>	Not enabled by default.
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Specify an IPv6 BGP peer and its AS number.	peer <i>ipv6-address</i> as-number <i>as-number</i>	Not configured by default.
5. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A
6. Enable the IPv6 MBGP peer.	peer <i>ipv6-address</i> enable	Not enabled by default.

Configuring a preferred value for routes from a peer or a peer group

If you both reference a routing policy and use the command **peer** { *ipv6-group-name* | *ipv6-address* } **preferred-value** *value* to set a preferred value for routes from a peer or a peer group, the routing policy sets the specified preferred value for the routes that match it. Other routes that do not match the routing policy use the value set through the command. If the preferred value in the routing policy is 0, the routes that match it will also use the value set through the **peer** { *ipv6-group-name* | *ipv6-address* } **preferred-value** *value* command. To learn how to use a routing policy to set a preferred value, see the **peer** { *ipv6-group-name* | *ipv6-address* } **route-policy** *route-policy-name* { **import** | **export** } command and the **apply preferred-value** *preferred-value* command. For more information about these commands, see *Layer 3—IP Routing Command Reference*.

To configure a preferred value for routes from a peer or a peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A
4. Specify a preferred value for routes received from the IPv6 MBGP peer or the peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } preferred-value <i>value</i>	Optional. The preferred value defaults to 0.

Controlling route distribution and reception

Configuration prerequisites

Before you configure this task, complete the following tasks:

- Enable IPv6.
- Configure basic IPv6 MBGP functions.

Injecting a local IPv6 MBGP route

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A
4. Inject a network into the IPv6 MBGP routing table.	network <i>ipv6-address prefix-length</i> [route-policy <i>route-policy-name</i> short-cut]	Not injected by default.

Configuring IPv6 MBGP route redistribution

Step	Command	Description
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 MBGP multicast address family view.	ipv6-family multicast	N/A

Step	Command	Description
4. Enable default route redistribution into the IPv6 MBGP routing table.	default-route imported	Optional. By default, default route redistribution is not allowed. If the default-route imported command is not configured, using the import-route command cannot redistribute any IGP default route.
5. Enable route redistribution from another routing protocol.	import-route <i>protocol</i> [<i>process-id</i> [med <i>med-value</i> route-policy <i>route-policy-name</i>] *]	Not enabled by default.

Configuring IPv6 MBGP route summarization

To reduce the routing table size on medium and large BGP networks, you must configure route summarization on IPv6 MBGP routers. BGP supports only manual summarization of IPv6 multicast routes.

To configure IPv6 MBGP route summarization:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A
4. Configure manual route summarization.	aggregate <i>ipv6-address prefix-length</i> [as-set attribute-policy <i>route-policy-name</i> detail-suppressed origin-policy <i>route-policy-name</i> suppress-policy <i>route-policy-name</i>]*	Not configured by default.

Advertising a default route to a peer or peer group

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A
4. Advertise a default route to an IPv6 MBGP peer or peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } default-route-advertise [route-policy <i>route-policy-name</i>]	Not advertised by default. With the peer default-route-advertise command executed, the router sends a default route with the next hop as itself to the specified IPv6 MBGP peer or the specified peer group, regardless of whether the default route is available in the routing table.

Configuring outbound IPv6 MBGP route filtering

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A
3. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A
4. Configure outbound IPv6 MBGP route filtering.	<ul style="list-style-type: none"> Configure the filtering of outgoing routes: filter-policy { acl6-number ipv6-prefix ipv6-prefix-name } export [protocol process-id] Specify an IPv6 ACL to filter routes advertised to a peer or a peer group: peer { ipv6-group-name ipv6-address } filter-policy acl6-number export Specify an AS path ACL to filter IPv6 MBGP routing information advertised to a peer or a peer group: peer { ipv6-group-name ipv6-address } as-path-acl as-path-acl-number export Specify an IPv6 prefix list to filter routes advertised to a peer or a peer group: peer { ipv6-group-name ipv6-address } ipv6-prefix ipv6-prefix-name export Apply a routing policy to routes advertised to a peer or a peer group: peer { ipv6-group-name ipv6-address } route-policy route-policy-name export 	<p>Use any of the commands. No filtering is configured by default.</p> <p>You can configure filter policies as required. If you configure multiple filter policies, they are applied in the following order:</p> <ol style="list-style-type: none"> filter-policy export peer filter-policy export peer as-path-acl export peer ipv6-prefix export peer route-policy export <p>A filter policy can be applied only after the previous one is passed. Routing information can be advertised only after passing all the configured filter policies.</p>

NOTE:

- Members of an IPv6 MBGP peer group must have the same outbound route filtering policy as the peer group.
- IPv6 BGP advertises the redistributed routes that pass the specified policy to the IPv6 MBGP peer.

Configuring inbound IPv6 MBGP route filtering

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A
3. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A

Step	Command	Remarks
4. Configure IPv6 MBGP inbound route filtering.	<ul style="list-style-type: none"> Configure inbound route filtering: filter-policy { <i>acl6-number</i> ipv6-prefix <i>ipv6-prefix-name</i> } import Apply a routing policy to routes from a peer or a peer group: peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } route-policy <i>route-policy-name</i> import Specify an IPv6 ACL to filter routes from a peer or a peer group: peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } filter-policy <i>acl6-number</i> import Specify an AS path ACL to filter IPv6 BGP routing information from a peer or a peer group: peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } as-path-acl <i>as-path-acl-number</i> import Specify an IPv6 prefix list to filter routes from a peer or a peer group: peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } ipv6-prefix <i>ipv6-prefix-name</i> import 	<p>Use any of the commands</p> <p>By default, advertised routes are not filtered.</p> <p>You can configure a filtering policy as required.</p> <p>If you configure several filtering policies, they are applied in the following sequence:</p> <ol style="list-style-type: none"> filter-policy import peer filter-policy import peer as-path-acl import peer ip-prefix import peer route-policy import <p>A filter policy can be applied only after the previous one is passed. Routing information can be received only after passing all the configured filter policies.</p>
10. Specify the upper limit of prefixes that can be imported from a peer or a peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } route-limit <i>limit</i> [<i>percentage</i>]	<p>Optional.</p> <p>The number is unlimited by default.</p>

NOTE:

A peer can have an inbound route filtering policy that is different from the policy of the peer group that it belongs to. That is, peer group members can have different inbound route filtering policies.

Configuring IPv6 MBGP route dampening

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A
4. Configure IPv6 MBGP route dampening parameters.	dampening [<i>half-life-reachable</i> <i>half-life-unreachable</i> <i>reuse</i> <i>suppress ceiling</i> route-policy <i>route-policy-name</i>]*	<p>Optional.</p> <p>Not configured by default.</p>

Configuring IPv6 MBGP route attributes

This section describes how to use IPv6 MBGP route attributes to affect IPv6 MBGP route selection. IPv6 MBGP route attributes involve:

- IPv6 MBGP protocol preference
- Default LOCAL_PREF attribute
- MED attribute
- NEXT_HOP attribute
- AS_PATH attribute

Configuration prerequisites

Before you configure IPv6 MBGP route attributes, complete the following tasks:

- Enable IPv6.
- Configure basic IPv6 MBGP functions.

Configuring IPv6 MBGP route preferences

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A
4. Configure preferences for external, internal, and local IPv6 MBGP routes.	preference { <i>external-preference</i> <i>internal-preference</i> <i>local-preference</i> route-policy <i>route-policy-name</i> }	Optional. The default preference values of external, internal, and local routes are 255, 255, and 130, respectively.

Configuring the default local preference

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A
4. Set the default local preference.	default local-preference <i>value</i>	Optional. By default, the default local preference is 100.

Configuring the MED attribute

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A

Step	Command	Remarks
3. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A
4. Configure a default MED value.	default med <i>med-value</i>	Optional. By default, the default <i>med-value</i> is 0.
5. Enable the comparison of the MED for routes from different ASs.	compare-different-as-med	Optional. Not enabled by default.
6. Enable the comparison of the MED for routes from each AS.	bestroute compare-med	Optional. Disabled by default.
7. Enable the comparison of the MED for routes from confederation peers.	bestroute med-confederation	Optional. Disabled by default.

Configuring the next hop attribute

You can use the **peer next-hop-local** command to specify the local router as the next hop of routes sent to an IPv6 multicast IBGP peer or a peer group. If load balancing is configured, the router specifies itself as the next hop of routes sent to the IPv6 multicast IBGP peer or the peer group regardless of whether the **peer next-hop-local** command is configured.

In a third-party next-hop network, that is, the local router has two IPv6 multicast EBGP peers in a broadcast network, the router does not specify itself as the next hop of routes sent to the EBGP peers by default.

To specify the router as the next hop of routes sent to a peer or a peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A
4. Configure the router as the next hop of routes sent to the peer or the peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } next-hop-local	Optional. By default, IPv6 MBGP specifies the local router as the next hop for routes sent to an EBGP peer or a peer group, but not for routes sent to an IBGP peer or a peer group.

Configuring the as_path attribute

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A

Step	Command	Remarks
4. Allow the local AS number to appear in the as-path of routes from a peer or a peer group and specify the number of times that the local AS number can appear in the as-path of routes from the peer or the peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } allow-as-loop [<i>number</i>]	Optional. Not allowed by default.
5. Disable IPv6 MBGP from considering the as_path during best route selection.	bestroute as-path-neglect	Optional. Enabled by default.
6. Configure updates to a peer or a peer group to carry only the public AS number.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } public-as-only	Optional. By default, outbound IPv6 MBGP updates can carry private AS numbers.

Tuning and optimizing IPv6 MBGP networks

Configuration prerequisites

Before you tune and optimize an OSPF network, complete the following tasks:

- Enable IPv6.
- Configure basic IPv6 MBGP functions.

Configuring IPv6 MBGP soft reset

After you modify a route selection policy, you must reset IPv6 MBGP connections to make the new one take effect.

The current IPv6 MBGP implementation supports the route-refresh feature that enables dynamic route refresh without terminating IPv6 MBGP connections.

If a peer that does not support route refresh exists in the network, you must configure the **peer keep-all-routes** command to save all routes from the peer. When the routing policy is changed, the system will update the IPv6 MBGP routing table and apply the new policy.

Performing soft reset through route refresh

If the peer is enabled with route refresh, when the IPv6 MBGP route selection policy is modified on a router, the router advertises a route-refresh message to its IPv6 MBGP peers, which resend their routing information to the router after they receive the message. Therefore, the local router can perform dynamic route update and apply the new policy without terminating IPv6 MBGP connections.

To configure IPv6 MBGP soft reset through route refresh:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A

Step	Command	Remarks
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Enable IPv6 BGP route refresh for a peer or a peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } capability-advertise route-refresh	Optional. Enabled by default.

Performing soft reset manually

If the peer does not support route refresh, you can use the **peer keep-all-routes** command to save all the route updates from the peer, and then use the **refresh bgp ipv6 multicast** command to soft-reset IPv6 MBGP connections to refresh the IPv6 MBGP routing table and apply the new policy without terminating IPv6 MBGP connections.

To perform soft reset manually:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A
5. Keep all routes from a peer or a peer group regardless of whether they pass the inbound filtering policy.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } keep-all-routes	Not kept by default.
6. Perform soft reset manually.	refresh bgp ipv6 multicast { all <i>ipv6-address</i> group <i>ipv6-group-name</i> external internal } { export import }	Optional.

Enabling the IPv6 MBGP orf capability

The BGP Outbound Route Filter (ORF) feature enables a BGP speaker to send a set of ORFs to its BGP peer through route-refresh messages. The peer then applies the ORFs, in addition to its local routing policies (if any), to filter updates to the BGP speaker, thus reducing the number of exchanged update messages and saving network resources.

After you enable the ORF capability, the local BGP router negotiates the ORF capability with the BGP peer through open messages. That is, the router determines whether to carry ORF information in messages, and if yes, whether to carry non-standard ORF information in the packets. After completing the negotiation process and establishing the neighboring relationship, the BGP router and its BGP peer can exchange ORF information through specific route-refresh messages.

For the parameters configured on both sides for ORF capability negotiation, see [Table 12](#).

To enable the IPv6 MBGP ORF capability:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A

Step	Command	Remarks
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Enable BGP route refresh for a peer or a peer group.	peer { group-name ipv6-address } capability-advertise route-refresh	Optional. Enabled by default. If this feature is not enabled, you must configure this command.
5. Enable the non-standard ORF capability for a BGP peer or a peer group.	peer { group-name ipv6-address } capability-advertise orf non-standard	Optional. By default, standard BGP ORF capability defined in RFC 5291 and RFC 5292 is supported. If this feature is not enabled, you must configure this command.
6. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A
7. Enable the ORF IP prefix negotiation capability for a BGP peer or a peer group.	peer { group-name ipv6-address } capability-advertise orf ip-prefix { both receive send }	Not enabled by default.

Table 12 Description of the send, receive, and both parameters and the negotiation result

Local parameter	Peer parameter	Negotiation result
send	<ul style="list-style-type: none"> • receive • both 	The ORF sending capability is enabled locally and the ORF receiving capability is enabled on the peer.
receive	<ul style="list-style-type: none"> • send • both 	The ORF receiving capability is enabled locally and the ORF sending capability is enabled on the peer.
both	both	Both the ORF sending and receiving capabilities are enabled locally and on the peer, respectively.

Configuring the maximum number of equal-cost routes for load-balancing

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp as-number	N/A
3. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A
4. Configure the maximum number of equal-cost routes for load balancing.	balance number	By default, load balancing is disabled.

Configuring a large scale IPv6 MBGP network

Configuration prerequisites

Before you configure the following tasks, you must configure basic IPv6 MBGP functions.

Configuring an IPv6 MBGP peer group

For easy management and configuration, you can organize some IPv6 MBGP peers that have the same route update policy into a peer group. A policy configured for a peer group applies to all the members in the group.

To create an IPv6 MBGP peer group, you must enable an existing IPv6 unicast peer group in IPv6 MBGP address family view.

Before adding an IPv6 MBGP peer to the IPv6 MBGP peer group, you must add the corresponding IPv6 BGP unicast peer to the corresponding IPv6 BGP unicast peer group.

To configure an IPv6 MBGP peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 address family view.	ipv6-family	N/A
4. Create an IPv6 BGP peer group.	group <i>ipv6-group-name</i> [external internal]	N/A
5. Add a peer to the peer group.	peer <i>ipv6-address</i> group <i>ipv6-group-name</i> [as-number <i>as-number</i>]	By default, no peer is added.
6. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A
7. Enable the configured IPv6 unicast BGP peer group to create the IPv6 MBGP peer group.	peer <i>ipv6-group-name</i> enable	N/A
8. Add the IPv6 MBGP peer into the peer group.	peer <i>ipv6-address</i> group <i>ipv6-group-name</i>	By default, no peer is added.

Configuring IPv6 MBGP community

A peer group enables a group of peers to share the same policy, and a community enables a group of IPv6 MBGP routers in multiple ASs to share the same policy. The community attribute is propagated among IPv6 MBGP peers and not restricted to AS boundaries.

You can reference a routing policy to modify the community attribute for routes sent to a peer. In addition, you can define extended community attributes as required.

For more information about routing policy configuration, see *Layer 3—IP Routing Configuration Guide*.

To advertise the community attribute to an IPv6 MBGP peer or a peer group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A
4. Advertise the community attribute to an IPv6 MBGP peer or a peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } advertise-community	By default, no community attribute is advertised to any peer group/peer.
5. Advertise the extended community attribute to an IPv6 MBGP peer or a peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } advertise-ext-community	By default, no extended community attribute is advertised to any peer or peer group.
6. Apply a routing policy to routes sent to an IPv6 MBGP peer or a peer group.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } route-policy <i>route-policy-name</i> export	Not configured by default.

NOTE:

You must configure a routing policy to define the community attribute, and apply the policy to outgoing routes.

Configuring an IPv6 MBGP route reflector

To guarantee connectivity between IPv6 multicast IBGP peers, you must make them fully meshed. However, this becomes unpractical when too many IPv6 multicast IBGP peers exist. Using route reflectors can solve the problem.

The clients of a route reflector should not be fully meshed, and the route reflector reflects the routes of a client to the other clients. If the clients are fully meshed, you must disable route reflection between clients to reduce routing costs.

If a cluster has multiple route reflectors, you must specify the same cluster ID for these route reflectors to avoid routing loops.

To configure an IPv6 BGP route reflector:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter BGP view.	bgp <i>as-number</i>	N/A
3. Enter IPv6 MBGP address family view.	ipv6-family multicast	N/A
4. Configure the router as a route reflector and specify an IPv6 MBGP peer or a peer group as its client.	peer { <i>ipv6-group-name</i> <i>ipv6-address</i> } reflect-client	Not configured by default.
5. Enable route reflection between clients.	reflect between-clients	Optional. Enabled by default.

Step	Command	Remarks
6. Configure the cluster ID of the route reflector.	reflector cluster-id <i>cluster-id</i>	Optional. By default, a route reflector uses its router ID as the cluster ID.

Displaying and maintaining IPv6 MBGP

Displaying IPv6 MBGP

Task	Command	Remarks
Display IPv6 MBGP peer group information.	display bgp ipv6 multicast group [<i>ipv6-group-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 MBGP routing information injected with the network command.	display bgp ipv6 multicast network [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the IPv6 MBGP AS path information of routes.	display bgp ipv6 multicast paths [<i>as-regular-expression</i> { begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 MBGP peer information or peer group information.	display bgp ipv6 multicast peer [[<i>ipv6-address</i>] verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the prefix entries in the ORF information of the specified BGP peer.	display bgp ipv6 multicast peer <i>ipv6-address</i> received ipv6-prefix [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 MBGP routing table information.	display bgp ipv6 multicast routing-table [<i>ipv6-address prefix-length</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 MBGP routing information that matches an AS path ACL.	display bgp ipv6 multicast routing-table as-path-acl <i>as-path-acl-number</i> [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 MBGP routing information with the specified community attribute.	display bgp ipv6 multicast routing-table community [<i>aa:nn<1-13></i>] [no-advertise no-export no-export-subconfed]* [whole-match] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display routing information matching an IPv6 MBGP community list.	display bgp ipv6 multicast routing-table community-list { { <i>basic-community-list-number</i> <i>comm-list-name</i> } [whole-match] <i>adv-community-list-number</i> } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 MBGP dampened routing information.	display bgp ipv6 multicast routing-table dampened [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 MBGP dampening parameter information.	display bgp ipv6 multicast routing-table dampening parameter [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 MBGP routing information originated from different ASs.	display bgp ipv6 multicast routing-table different-origin-as [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display IPv6 MBGP routing flap statistics.	display bgp ipv6 multicast routing-table flap-info [regular-expression <i>as-regular-expression</i> [as-path-acl <i>as-path-acl-number</i> <i>ipv6-address</i> <i>prefix-length</i> [longer-match]] [{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display the IPv6 MBGP routes received from or advertised to the IPv6 MBGP peer or peer group.	display bgp ipv6 multicast routing-table peer <i>ipv6-address</i> { advertised-routes received-routes } [<i>network-address</i> <i>prefix-length</i> statistic] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display IPv6 multicast routing information matching an AS regular expression.	display bgp ipv6 multicast routing-table regular-expression <i>as-regular-expression</i>	Available in any view
Display IPv6 MBGP routing statistics.	display bgp ipv6 multicast routing-table statistic [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the IPv6 MBGP routing table information.	display ipv6 multicast routing-table [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the multicast routing information of the specified destination address.	display ipv6 multicast routing-table <i>ipv6-address</i> <i>prefix-length</i> [longer-match] [verbose] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Resetting IPv6 MBGP connections

When you change an IPv6 MBGP routing policy, you can make the new configuration effective by resetting the IPv6 MBGP connections.

Task	Command	Remarks
Reset the specified IPv6 MBGP connections.	reset bgp ipv6 multicast { <i>as-number</i> <i>ipv6-address</i> all group <i>ipv6-group-name</i> external internal }	Available in user view

Clearing IPv6 MBGP information

Task	Command	Remarks
Clear dampened IPv6 MBGP routing information and release suppressed routes.	reset bgp ipv6 multicast dampening [<i>ipv6-address</i> <i>prefix-length</i>]	Available in user view
Clear IPv6 MBGP route flap statistics.	reset bgp ipv6 multicast flap-info [<i>ipv6-address/prefix-length</i> regex <i>as-path-regex</i> as-path-acl <i>as-path-acl-number</i>]	Available in user view

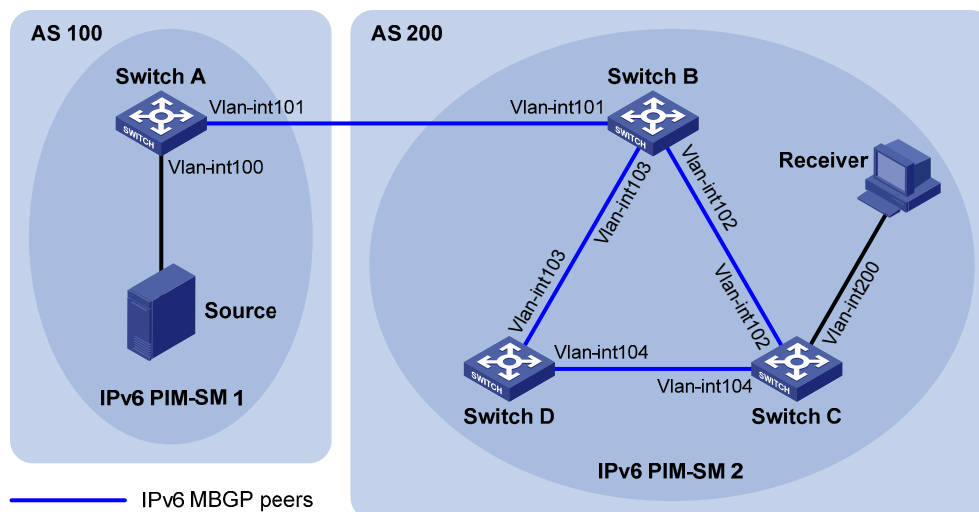
IPv6 MBGP configuration example

Network requirements

As shown in the following figure:

- IPv6 PIM-SM 1 is in AS 100, and IPv6 PIM-SM 2 is in AS 200. OSPFv3 is the IGP in the two ASs, and IPv6 MBGP runs between the two ASs to exchange IPv6 multicast route information.
- The IPv6 multicast source belongs to IPv6 PIM-SM 1 and the receiver belongs to IPv6 PIM-SM 2.
- The VLAN-interface 101 of Switch A and Switch B must be configured as the C-BSR and C-RP of the IPv6 PIM-SM domains, respectively.

Figure 108 Network diagram



Device	Interface	IP address	Device	Interface	IP address
Source	N/A	1002::100/64	Switch C	Vlan-int200	3002::1/64
Switch A	Vlan-int100	1002::1/64	Switch B	Vlan-int102	2001::2/64
	Vlan-int101	1001::1/64		Vlan-int104	3001::1/64
Switch B	Vlan-int101	1001::2/64	Switch D	Vlan-int103	2002::2/64
	Vlan-int102	2001::1/64		Vlan-int104	3001::2/64
	Vlan-int103	2002::1/64			

Configuration procedure

1. Configure IPv6 addresses for interfaces as shown in Figure 108. (Details not shown.)
2. Configure OSPFv3. (Details not shown.)
3. Enable IPv6 multicast routing, IPv6 PIM-SM and MLD, and configure an IPv6 PIM-SM domain border:

Enable IPv6 multicast routing on Switch A, and enable IPv6 PIM-SM on each interface.

```
<SwitchA> system-view
[SwitchA] multicast ipv6 routing-enable
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] pim ipv6 sm
[SwitchA-Vlan-interface100] quit
[SwitchA] interface vlan-interface 101
```

```
[SwitchA-Vlan-interface101] pim ipv6 sm
[SwitchA-Vlan-interface101] quit
```

The configuration on Switch B and Switch D is similar to the configuration on Switch A.

Enable IPv6 multicast routing on Switch C, enable IPv6 PIM-SM on each interface, and enable MLD on the host-side interface VLAN-interface 200.

```
<SwitchC> system-view
[SwitchC] multicast ipv6 routing-enable
[SwitchC] interface vlan-interface 102
[SwitchC-Vlan-interface102] pim ipv6 sm
[SwitchC-Vlan-interface102] quit
[SwitchC] interface vlan-interface 104
[SwitchC-Vlan-interface104] pim ipv6 sm
[SwitchC-Vlan-interface104] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] pim ipv6 sm
[SwitchC-Vlan-interface200] mld enable
[SwitchC-Vlan-interface200] quit
```

Configure an IPv6 PIM domain border on Switch A.

```
[SwitchA] interface vlan-interface 101
[SwitchA-Vlan-interface101] pim ipv6 bsr-boundary
[SwitchA-Vlan-interface101] quit
```

Configure an IPv6 PIM domain border on Switch B.

```
[SwitchB] interface vlan-interface 101
[SwitchB-Vlan-interface101] pim ipv6 bsr-boundary
[SwitchB-Vlan-interface101] quit
```

4. Configure the position of C-BSR and C-RP:

Configure the position of C-BSR and C-RP on Switch A.

```
[SwitchA] pim ipv6
[SwitchA-pim6] c-bsr 1001::1
[SwitchA-pim6] c-rp 1001::1
[SwitchA-pim6] quit
```

Configure the position of C-BSR and C-RP on Switch B.

```
[SwitchB] pim ipv6
[SwitchB-pim6] c-bsr 1001::2
[SwitchB-pim6] c-rp 1001::2
[SwitchB-pim6] quit
```

5. Configure BGP, specify the IPv6 MBGP peer and enable direct route redistribution:

On Switch A, configure the IPv6 MBGP peer and enable direct route redistribution.

```
[SwitchA] ipv6
[SwitchA] bgp 100
[SwitchA-bgp] router-id 1.1.1.1
[SwitchA-bgp] ipv6-family
[SwitchA-bgp-af-ipv6] peer 1001::2 as-number 200
[SwitchA-bgp-af-ipv6] import-route direct
[SwitchA-bgp-af-ipv6] quit
[SwitchA-bgp] ipv6-family multicast
```

```
[SwitchA-bgp-af-ipv6-mul] peer 1001::2 enable
[SwitchA-bgp-af-ipv6-mul] import-route direct
[SwitchA-bgp-af-ipv6-mul] quit
[SwitchA-bgp] quit
```

On Switch B, configure the IPv6 MBGP peers and redistribute OSPF routes.

```
[SwitchB] ipv6
[SwitchB] bgp 200
[SwitchB-bgp] router-id 2.2.2.2
[SwitchB-bgp] ipv6-family
[SwitchB-bgp-af-ipv6] peer 1001::1 as-number 100
[SwitchB-bgp-af-ipv6] import-route ospfv3 1
[SwitchB-bgp-af-ipv6] quit
[SwitchB-bgp] ipv6-family multicast
[SwitchB-bgp-af-ipv6-mul] peer 1001::1 enable
[SwitchB-bgp-af-ipv6-mul] import-route ospfv3 1
[SwitchB-bgp-af-ipv6-mul] quit
[SwitchB-bgp] quit
```

6. Verify the configuration:

You can use the **display bgp ipv6 multicast peer** command to display IPv6 MBGP peers on a switch. For example, display IPv6 MBGP peers on Switch B.

```
[SwitchB] display bgp ipv6 multicast peer
```

```
BGP local router ID : 2.2.2.2
Local AS number : 200
Total number of peers : 3                Peers in established state : 3
```

Peer	AS	MsgRcvd	MsgSent	OutQ	PrefRcv	Up/Down	State
1001::1	100	56	56	0	0	00:40:54	Established

Index

A C D E I M O P T

A

Adjusting IGMP performance, [98](#)

Adjusting MLD performance, [319](#)

Appendix, [279](#)

Appendix, [51](#)

C

Configuration examples, [82](#)

Configuration task list, [76](#)

Configuration task list, [302](#)

Configuring a large scale IPv6 MBGP network, [425](#)

Configuring a large scale MBGP network, [233](#)

Configuring a port-based IPv6 multicast VLAN, [290](#)

Configuring a port-based multicast VLAN, [62](#)

Configuring a sub-VLAN-based IPv6 multicast VLAN, [289](#)

Configuring a sub-VLAN-based multicast VLAN, [61](#)

Configuring an IGMP snooping policy, [28](#)

Configuring an MLD snooping policy, [256](#)

Configuring an MSDP peer connection, [197](#)

Configuring basic IGMP functions, [95](#)

Configuring basic IGMP snooping functions, [19](#)

Configuring basic IPv6 MBGP functions, [415](#)

Configuring basic MBGP functions, [222](#)

Configuring basic MLD functions, [316](#)

Configuring basic MLD snooping functions, [247](#)

Configuring basic MSDP functions, [195](#)

Configuring BIDIR-PIM, [147](#)

Configuring IGMP proxying, [104](#)

Configuring IGMP snooping port functions, [21](#)

Configuring IGMP snooping proxying, [27](#)

Configuring IGMP snooping querier, [25](#)

Configuring IGMP SSM mapping, [103](#)

Configuring IPv6 BIDIR-PIM, [366](#)

Configuring IPv6 MBGP route attributes, [419](#)

Configuring IPv6 multicast routing and forwarding, [303](#)

Configuring IPv6 PIM common features, [375](#)

Configuring IPv6 PIM snooping, [282](#)

Configuring IPv6 PIM-DM, [353](#)

Configuring IPv6 PIM-SM, [356](#)

Configuring IPv6 PIM-SSM, [373](#)

Configuring MBGP route attributes, [228](#)

Configuring MLD proxying, [325](#)

Configuring MLD snooping port functions, [249](#)

Configuring MLD snooping proxying, [255](#)

Configuring MLD snooping querier, [253](#)

Configuring MLD SSM mapping, [324](#)

Configuring multicast routing and forwarding, [77](#)

Configuring PIM common features, [156](#)

Configuring PIM snooping, [54](#)

Configuring PIM-DM, [132](#)

Configuring PIM-SM, [135](#)

Configuring PIM-SSM, [154](#)

Configuring SA messages related parameters, [199](#)

Controlling route advertisement and reception, [223](#)

Controlling route distribution and reception, [416](#)

D

Displaying and maintaining IGMP, [105](#)

Displaying and maintaining IGMP snooping, [35](#)

Displaying and maintaining IPv6 MBGP, [427](#)

Displaying and maintaining IPv6 multicast routing and forwarding, [305](#)

Displaying and maintaining IPv6 multicast VLAN, [292](#)

Displaying and maintaining IPv6 PIM, [381](#)

Displaying and maintaining IPv6 PIM snooping, [283](#)

Displaying and maintaining MBGP, [235](#)

Displaying and maintaining MLD, [326](#)

Displaying and maintaining MLD snooping, [263](#)

Displaying and maintaining MSDP, [202](#)

Displaying and maintaining multicast routing and forwarding, [81](#)

Displaying and maintaining multicast VLAN, [64](#)

Displaying and maintaining PIM, [163](#)

Displaying and maintaining PIM snooping, [55](#)

E

Enabling IP multicast routing, [77](#)
Enabling IPv6 multicast routing, [303](#)

I

IGMP configuration examples, [108](#)
IGMP configuration task list, [94](#)
IGMP snooping configuration examples, [36](#)
IGMP snooping configuration task list, [18](#)
Introduction to multicast, [1](#)
IPv6 MBGP configuration example, [429](#)
IPv6 MBGP overview, [414](#)
IPv6 multicast VLAN configuration examples, [292](#)
IPv6 multicast VLAN configuration task list, [289](#)
IPv6 PIM configuration examples, [383](#)
IPv6 PIM snooping configuration example, [283](#)

M

MBGP configuration example, [237](#)
MBGP configuration task list, [221](#)
MBGP overview, [221](#)
MLD configuration examples, [328](#)
MLD configuration task list, [316](#)
MLD snooping configuration examples, [264](#)
MLD snooping configuration task list, [246](#)
MSDP configuration examples, [203](#)
MSDP configuration task list, [194](#)
Multicast architecture, [5](#)
Multicast models, [5](#)
Multicast packet forwarding mechanism, [11](#)
Multicast support for VPNs, [11](#)
Multicast VLAN configuration examples, [64](#)
Multicast VLAN configuration task list, [61](#)

O

Overview, [300](#)

Overview, [241](#)
Overview, [281](#)
Overview, [287](#)
Overview, [337](#)
Overview, [13](#)
Overview, [88](#)
Overview, [53](#)
Overview, [59](#)
Overview, [308](#)
Overview, [72](#)
Overview, [189](#)

P

PIM configuration examples, [165](#)
PIM overview, [116](#)
PIM snooping configuration example, [55](#)
Protocols and standards, [221](#)

T

Troubleshooting IGMP, [114](#)
Troubleshooting IGMP snooping, [51](#)
Troubleshooting IPv6 multicast policy configuration, [306](#)
Troubleshooting IPv6 PIM configuration, [410](#)
Troubleshooting IPv6 PIM snooping, [286](#)
Troubleshooting MLD, [335](#)
Troubleshooting MLD snooping, [279](#)
Troubleshooting MSDP, [219](#)
Troubleshooting multicast routing and forwarding, [86](#)
Troubleshooting PIM, [186](#)
Troubleshooting PIM snooping, [58](#)
Tuning and optimizing IPv6 MBGP networks, [422](#)
Tuning and optimizing MBGP networks, [230](#)

Contents

Configuring ACLs	1
Overview	1
Applications on the switch	1
ACL categories	1
Numbering and naming ACLs	2
Match order	2
ACL rule comments and rule range remarks	3
ACL rule numbering	3
Fragments filtering with ACLs	4
ACL configuration task list	4
Configuring a time range	4
Configuring a basic ACL	5
Configuring an IPv4 basic ACL	5
Configuring an IPv6 basic ACL	6
Configuring an advanced ACL	6
Configuring an IPv4 advanced ACL	6
Configuring an IPv6 advanced ACL	7
Configuring an Ethernet frame header ACL	9
Copying an ACL	9
Copying an IPv4 ACL	10
Copying an IPv6 ACL	10
Packet filtering with ACLs	10
Applying an IPv4 or Ethernet frame header ACL for packet filtering	10
Applying an IPv6 ACL for packet filtering	11
Displaying and maintaining ACLs	11
Configuration example of using ACL for device management	12
Network requirements	12
Configuration procedure	12
IPv4 packet filtering configuration example	13
Network requirements	13
Configuration procedure	14
IPv6 packet filtering configuration example	14
Network requirements	14
Configuration procedure	14
QoS overview	16
QoS service models	16
Best-effort service model	16
IntServ model	16
DiffServ model	16
QoS techniques	17
QoS configuration approaches	18
MQC approach	18
Non-MQC approach	18
Configuring a QoS policy	19
Overview	19
Defining a class	19
Configuration restrictions and guidelines	20

Configuration procedure	20
Defining a traffic behavior	21
Defining a policy	22
Configuration restrictions and guidelines	22
Configuration procedure	22
Applying the QoS policy	22
Applying the QoS policy to an interface	23
Applying the QoS policy to online users	23
Applying the QoS policy to a VLAN	24
Applying the QoS policy globally	24
Applying the QoS policy to the control plane	25
Displaying and maintaining QoS policies	25
Configuring priority mapping	27
Overview	27
Types of priorities	27
Priority mapping tables	27
Priority trust mode on a port	28
Priority mapping procedure	28
Configuration guidelines	29
Configuring a priority mapping table	29
Configuring a port to trust packet priority for priority mapping	30
Changing the port priority of an interface	30
Displaying priority mappings	31
Priority trust mode configuration example	31
Network requirements	31
Configuration procedure	32
Priority mapping table and priority marking configuration example	32
Network requirements	32
Configuration procedure	33
Configuring traffic policing, traffic shaping, and line rate	35
Overview	35
Traffic evaluation and token buckets	35
Traffic policing	36
Traffic shaping	37
Line rate	38
Configuring traffic policing	39
Configuration restrictions and guidelines	39
Configuration procedure	39
Configuring GTS	40
Configuring the line rate	40
Displaying and maintaining traffic policing, GTS, and line rate	41
Traffic policing configuration example	41
Network requirements	41
Configuration procedures	42
Configuring congestion management	44
Overview	44
Congestion management techniques	44
SP queuing	45
WRR queuing	45
WFQ queuing	47
SP+WRR queuing	47
SP+WFQ queuing	48
Configuring SP queuing	48

Configuration procedure	48
Configuration example	48
Configuring WRR queuing	49
Configuration procedure	49
Configuration example	49
Configuring WFQ queuing	50
Configuration procedure	50
Configuration example	51
Configuring SP+WRR queuing	52
Configuration procedure	52
Configuration example	52
Configuring SP+WFQ queuing	53
Configuration procedure	53
Configuration example	54
Configuring congestion avoidance (available only on the 5500 EI)	55
Overview	55
Tail drop	55
RED and WRED	55
Introduction to WRED configuration	56
Configuration procedure	56
Configuration examples	57
Displaying and maintaining WRED	57
Configuring traffic filtering	58
Configuration procedure	58
Traffic filtering configuration example	59
Network requirements	59
Configuration procedure	59
Configuring priority marking	60
Color-based priority marking	60
Coloring a packet	60
Marking packets based on their colors	60
Configuration procedure	61
Local precedence re-marking configuration example	62
Network requirements	62
Configuration procedure	62
Configuring traffic redirecting	65
Configuration restrictions and guidelines	65
Configuration procedure	65
Redirect-to-next hop configuration example	66
Network requirements	66
Configuration procedure	67
Configuring class-based accounting	69
Configuration procedure	69
Displaying and maintaining traffic accounting	69
Class-based accounting configuration example	70
Network requirements	70
Configuration procedure	70
Configuring the data buffer	72
Overview	72
Data buffer	72
Data buffer allocation	72

Data buffer configuration approaches.....	73
Using the burst function to configure the data buffer setup.....	74
Manually configuring the data buffer setup.....	74
Manually configuring the data buffer.....	74
Configuring the cell resource.....	74
Configuring the packet resource.....	76
Applying the data buffer settings.....	76
Appendix A Default priority mapping tables.....	77
Uncolored priority mapping tables.....	77
Appendix B Packet precedences.....	78
IP precedence and DSCP values.....	78
802.1p priority.....	79
Index.....	81

Configuring ACLs

- Unless otherwise stated, ACLs refer to both IPv4 and IPv6 ACLs throughout this document.
- The term *interface* in the routing features refers to VLAN interfaces, bridge mode (Layer 2) and route mode (Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*). 5500 SI Switch Series does not support Layer 3 Ethernet ports.

Overview

An access control list (ACL) is a set of rules (or permit or deny statements) for identifying traffic based on criteria such as source IP address, destination IP address, and port number.

ACLs are primarily used for packet filtering. A packet filter drops packets that match a deny rule and permits packets that match a permit rule. ACLs are also used by many modules, QoS and IP routing for example, for traffic classification and identification.

Applications on the switch

An ACL is implemented in hardware or software, depending on the module that uses it. If the module, the packet filter or QoS module for example, is implemented in hardware, the ACL is applied to hardware to process traffic. If the module, the routing or user interface access control module (Telnet, SNMP, or web) for example, is implemented in software, the ACL is applied to software to process traffic.

The user interface access control module denies packets that do not match any ACL. Some modules, QoS for example, ignore the permit or deny action in ACL rules and do not base their drop or forwarding decisions on the action set in ACL rules. See the specified module for information about ACL application.

ACL categories

Category	ACL number	IP version	Match criteria
Basic ACLs	2000 to 2999	IPv4	Source IPv4 address
		IPv6	Source IPv6 address
Advanced ACLs	3000 to 3999	IPv4	Source IPv4 address, destination IPv4 address, packet priority, protocols over IPv4, and other Layer 3 and Layer 4 header fields
		IPv6	Source IPv6 address, destination IPv6 address, packet priority, protocols over IPv6, and other Layer 3 and Layer 4 header fields
Ethernet frame header ACLs	4000 to 4999	IPv4 and IPv6	Layer 2 header fields, such as source and destination MAC addresses, 802.1p priority, and link layer protocol type

Numbering and naming ACLs

Each ACL category has a unique range of ACL numbers. When creating an ACL, you must assign it a number. In addition, you can assign the ACL a name for ease of identification. After creating an ACL with a name, you cannot rename it or delete its name.

For an Ethernet frame header ACL, the ACL number and name must be globally unique. For an IPv4 basic or advanced ACLs, its ACL number and name must be unique among all IPv4 ACLs, and for an IPv6 basic or advanced ACL, its ACL number and name must be unique among all IPv6 ACLs. You can assign an IPv4 ACL and an IPv6 ACL the same number and name.

Match order

The rules in an ACL are sorted in a specific order. When a packet matches a rule, the device stops the match process and performs the action defined in the rule. If an ACL contains overlapping or conflicting rules, the matching result and action to take depend on the rule order.

The following ACL match orders are available:

- **config**—Sorts ACL rules in ascending order of rule ID. A rule with a lower ID is matched before a rule with a higher ID. If you use this approach, carefully check the rules and their order.
- **auto**—Sorts ACL rules in depth-first order. Depth-first ordering guarantees that any subset of a rule is always matched before the rule. [Table 1](#) lists the sequence of tie breakers that depth-first ordering uses to sort rules for each type of ACL.

Table 1 Sort ACL rules in depth-first order

ACL category	Sequence of tie breakers
IPv4 basic ACL	<ol style="list-style-type: none">1. VPN instance2. More 0s in the source IP address wildcard (more 0s means a narrower IP address range)3. Rule configured earlier
IPv4 advanced ACL	<ol style="list-style-type: none">1. VPN instance2. Specific protocol type rather than IP (IP represents any protocol over IP)3. More 0s in the source IP address wildcard mask4. More 0s in the destination IP address wildcard5. Narrower TCP/UDP service port number range6. Rule configured earlier
IPv6 basic ACL	<ol style="list-style-type: none">1. VPN instance2. Longer prefix for the source IP address (a longer prefix means a narrower IP address range)3. Rule configured earlier
IPv6 advanced ACL	<ol style="list-style-type: none">1. VPN instance2. Specific protocol type rather than IP (IP represents any protocol over IPv6)3. Longer prefix for the source IPv6 address4. Longer prefix for the destination IPv6 address5. Narrower TCP/UDP service port number range6. Rule configured earlier

ACL category	Sequence of tie breakers
Ethernet frame header ACL	7. More 1s in the source MAC address mask (more 1s means a smaller MAC address)
	8. More 1s in the destination MAC address mask
	9. Rule configured earlier

A wildcard mask, also called an inverse mask, is a 32-bit binary and represented in dotted decimal notation. In contrast to a network mask, the 0 bits in a wildcard mask represent "do care" bits, and the 1 bits represent "don't care" bits. If the "do care" bits in an IP address are identical to the "do care" bits in an IP address criterion, the IP address matches the criterion. All "don't care" bits are ignored. The 0s and 1s in a wildcard mask can be noncontiguous. For example, 0.255.0.255 is a valid wildcard mask

NOTE:

Only 5500 EI Switch Series supports VPN instance configuration in an ACL rule. .

ACL rule comments and rule range remarks

You can add a comment about an ACL rule to make it easy to understand. The rule comment appears below the rule statement.

You can also add a rule range remark to indicate the start or end of a range of rules created for the same purpose. A rule range remark always appears above the specified ACL rule. If the specified rule has not been created yet, the position of the comment in the ACL is as follows:

- If the match order is config, the remark is inserted into the ACL in descending order of rule ID.
- If the match order is auto, the remark is placed at the end of the ACL. After you create the rule, the remark appears above the rule.

For more information about how to use rule range remarks, see the **rule remark** command in *ACL and QoS Command Reference* for your device.

ACL rule numbering

What is the ACL rule numbering step

If you do not assign an ID to the rule you are creating, the system automatically assigns it a rule ID. The rule numbering step sets the increment by which the system automatically numbers rules. For example, the default ACL rule numbering step is 5. If you do not assign IDs to rules you are creating, they are automatically numbered 0, 5, 10, 15, and so on. The wider the numbering step, the more rules you can insert between two rules.

By introducing a gap between rules rather than contiguously numbering rules, you have the flexibility of inserting rules in an ACL. This feature is important for a config order ACL, where ACL rules are matched in ascending order of rule ID.

Automatic rule numbering and renumbering

The ID automatically assigned to an ACL rule takes the nearest higher multiple of the numbering step to the current highest rule ID, starting with 0.

For example, if the numbering step is 5 (the default), and there are five ACL rules numbered 0, 5, 9, 10, and 12, the newly defined rule is numbered 15. If the ACL does not contain any rule, the first rule is numbered 0.

Whenever the step changes, the rules are renumbered, starting from 0. For example, if there are five rules numbered 5, 10, 13, 15, and 20, changing the step from 5 to 2 causes the rules to be renumbered 0, 2, 4, 6, and 8.

Fragments filtering with ACLs

Traditional packet filtering matches only first fragments of packets, and allows all subsequent non-first fragments to pass through. Attackers can fabricate non-first fragments to attack networks.

To avoid the risks, the HP ACL implementation:

- Filters all fragments by default, including non-first fragments.
- Allows for matching criteria modification, for example, filters non-first fragments only.

ACL configuration task list

Task	Remarks
Configuring a time range	Optional Applicable to IPv4 and IPv6 ACLs.
Configuring a basic ACL	Required
Configuring an advanced ACL	Configure at least one task.
Configuring an Ethernet frame header ACL	Applicable to IPv4 and IPv6 except that simple ACLs are for IPv6.
Copying an ACL	Optional Applicable to IPv4 and IPv6.
Packet filtering with ACLs	Optional Applicable to IPv4 and IPv6.

Configuring a time range

You can implement ACL rules based on the time of day by applying a time range to them. A time-based ACL rule only takes effect in any time periods specified by the time range.

The following basic types of time range are available:

- **Periodic time range**—Recurrs periodically on a day or days of the week.
- **Absolute time range**—Represents only a period of time and does not recur.

You can create a maximum of 256 time ranges, each with a maximum of 32 periodic statements and 12 absolute statements. The active period of a time range is calculated as follows:

1. Combining all periodic statements.
2. Combining all absolute statements.
3. Taking the intersection of the two statement sets as the active period of the time range.

To configure a time range:

Step	Command	Remarks
1. Enter system view.	<code>system-view</code>	N/A

Step	Command	Remarks
2. Configure a time range.	time-range <i>time-range-name</i> { <i>start-time to end-time days</i> [from <i>time1 date1</i>] [to <i>time2 date2</i>] from <i>time1 date1</i> [to <i>time2 date2</i>] to <i>time2 date2</i> }	By default, no time range exists. Repeat this command with the same time range name to create multiple statements for a time range.

Configuring a basic ACL

Configuring an IPv4 basic ACL

IPv4 basic ACLs match packets based only on source IP addresses.

To configure an IPv4 basic ACL:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create an IPv4 basic ACL and enter its view.	acl number <i>acl-number</i> [name <i>acl-name</i>] [match-order { auto config }]	By default, no ACL exists. IPv4 basic ACLs are numbered in the range of 2000 to 2999. You can use the acl name <i>acl-name</i> command to enter the view of a named IPv4 ACL.
3. Configure a description for the IPv4 basic ACL.	description <i>text</i>	Optional. By default, an IPv4 basic ACL has no ACL description.
4. Set the rule numbering step.	step <i>step-value</i>	Optional. The default setting is 5.
5. Create or edit a rule.	rule [<i>rule-id</i>] { deny permit } [counting fragment logging source { <i>sour-addr</i> <i>sour-wildcard</i> any } time-range <i>time-range-name</i> vpn-instance <i>vpn-instance-name</i>] *	By default, an IPv4 basic ACL does not contain any rule. The vpn-instance <i>vpn-instanced-name</i> option is not available on a 5500 SI switch. If the ACL is for QoS traffic classification or packet filtering, do not specify the vpn-instance keyword. This keyword can cause ACL application failure. The logging and counting keywords (even if specified) do not take effect for QoS policies.
6. Add or edit a rule comment.	rule <i>rule-id</i> comment <i>text</i>	Optional. By default, no rule comments are configured.
7. Add or edit a rule range remark.	rule [<i>rule-id</i>] remark <i>text</i>	Optional. By default, no rule range remarks are configured.
8. Enable counting ACL rule matches performed in hardware.	hardware-count enable	Optional. Disabled by default. When the ACL is referenced by a QoS policy, this command does not take effect.

Configuring an IPv6 basic ACL

To configure an IPv6 basic ACL:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create an IPv6 basic ACL view and enter its view.	acl ipv6 number <i>acl6-number</i> [name <i>acl6-name</i>] [match-order { auto config }]	By default, no ACL exists. IPv6 basic ACLs are numbered in the range of 2000 to 2999. You can use the acl ipv6 name <i>acl6-name</i> command to enter the view of a named IPv6 ACL.
3. Configure a description for the IPv6 basic ACL.	description <i>text</i>	Optional. By default, an IPv6 basic ACL has no ACL description.
4. Set the rule numbering step.	step <i>step-value</i>	Optional. The default setting is 5.
5. Create or edit a rule.	rule [<i>rule-id</i>] { deny permit } [counting fragment logging routing [type <i>routing-type</i>] source { <i>ipv6-address</i> <i>prefix-length</i> <i>ipv6-address/prefix-length</i> any } time-range <i>time-range-name</i> vpn-instance <i>vpn-instance-name</i>] *	By default, an IPv6 basic ACL does not contain any rule. The vpn-instance <i>vpn-instance-name</i> option is not available on a 5500 SI switch. If the ACL is for QoS traffic classification or packet filtering, do not specify the fragment , routing , and vpn-instance keywords. The keywords can cause ACL application failure. The logging and counting keywords (even if specified) do not take effect for QoS.
6. Add or edit a rule comment.	rule <i>rule-id</i> comment <i>text</i>	Optional. By default, no rule comments are configured.
7. Add or edit a rule range remark.	rule [<i>rule-id</i>] remark <i>text</i>	Optional. By default, no rule range remarks are configured.
8. Enable counting ACL rule matches performed in hardware.	hardware-count enable	Optional. Disabled by default. When the ACL is referenced by a QoS policy, this command does not take effect.

Configuring an advanced ACL

Configuring an IPv4 advanced ACL

IPv4 advanced ACLs match packets based on source IP addresses, destination IP addresses, packet priorities, protocols over IP, and other protocol header information, such as TCP/UDP source and destination port numbers, TCP flags, ICMP message types, and ICMP message codes.

Compared to IPv4 basic ACLs, IPv4 advanced ACLs allow more flexible and accurate filtering.

To configure an IPv4 advanced ACL:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create an IPv4 advanced ACL and enter its view.	acl number <i>acl-number</i> [name <i>acl-name</i>] [match-order { auto config }]	By default, no ACL exists. IPv4 advanced ACLs are numbered in the range of 3000 to 3999. You can use the acl name <i>acl-name</i> command to enter the view of a named IPv4 ACL.
3. Configure a description for the IPv4 advanced ACL.	description <i>text</i>	Optional. By default, an IPv4 advanced ACL has no ACL description.
4. Set the rule numbering step.	step <i>step-value</i>	Optional. The default setting is 5.
5. Create or edit a rule.	rule [<i>rule-id</i>] { deny permit } protocol [{ { ack <i>ack-value</i> fin <i>fin-value</i> psh <i>psh-value</i> rst <i>rst-value</i> syn <i>syn-value</i> urg <i>urg-value</i> } * established } counting destination { <i>dest-addr</i> <i>dest-wildcard</i> any } destination-port <i>operator</i> <i>port1</i> [<i>port2</i>] dscp <i>dscp</i> fragment icmp-type { <i>icmp-type</i> [<i>icmp-code</i>] <i>icmp-message</i> } logging precedence <i>precedence</i> source { <i>sour-addr</i> <i>sour-wildcard</i> any } source-port <i>operator</i> <i>port1</i> [<i>port2</i>] time-range <i>time-range-name</i> tos <i>tos</i> vpn-instance <i>vpn-instance-name</i>] *	By default, an IPv4 advanced ACL does not contain any rule. The vpn-instance <i>vpn-instance-name</i> option is not available on a 5500 SI switch. If an IPv4 advanced ACL is for QoS traffic classification or packet filtering, do not specify the vpn-instance keyword or specify neg for the <i>operator</i> argument. The logging and counting keywords (even if specified) do not take effect for QoS traffic classification.
6. Add or edit a rule comment.	rule <i>rule-id</i> comment <i>text</i>	Optional. By default, no rule comments are configured.
7. Add or edit a rule range remark.	rule [<i>rule-id</i>] remark <i>text</i>	Optional. By default, no rule range remarks are configured.
8. Enable counting ACL rule matches performed in hardware.	hardware-count enable	Optional. Disabled by default. When the ACL is referenced by a QoS policy, this command does not take effect.

Configuring an IPv6 advanced ACL

IPv6 advanced ACLs match packets based on the source IPv6 addresses, destination IPv6 addresses, packet priorities, protocols carried over IPv6, and other protocol header fields such as the TCP/UDP source port number, TCP/UDP destination port number, ICMPv6 message type, and ICMPv6 message code.

Compared to IPv6 basic ACLs, IPv6 advanced ACLs allow more flexible and accurate filtering.

To configure an IPv6 advanced ACL:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create an IPv6 advanced ACL and enter its view.	acl ipv6 number <i>acl6-number</i> [name <i>acl6-name</i>] [match-order { auto config }]	By default, no ACL exists. IPv6 advanced ACLs are numbered in the range of 3000 to 3999. You can use the acl ipv6 name <i>acl6-name</i> command to enter the view of a named IPv6 ACL.
3. Configure a description for the IPv6 advanced ACL.	description <i>text</i>	Optional. By default, an IPv6 advanced ACL has no ACL description.
4. Set the rule numbering step.	step <i>step-value</i>	Optional. 5 by default.
5. Create or edit a rule.	rule [<i>rule-id</i>] { deny permit } <i>protocol</i> [{ { ack <i>ack-value</i> fin <i>fin-value</i> psh <i>psh-value</i> rst <i>rst-value</i> syn <i>syn-value</i> urg <i>urg-value</i> } * established } counting destination { <i>dest</i> <i>dest-prefix</i> <i>dest/dest-prefix</i> any } destination-port <i>operator port1</i> [<i>port2</i>] dscp <i>dscp</i> flow-label <i>flow-label-value</i> fragment icmp6-type { <i>icmp6-type</i> <i>icmp6-code</i> <i>icmp6-message</i> } logging routing [<i>type</i> <i>routing-type</i>] source { <i>source</i> <i>source-prefix</i> <i>source/source-prefix</i> any } source-port <i>operator port1</i> [<i>port2</i>] time-range <i>time-range-name</i> vpn-instance <i>vpn-instance-name</i>] *	By default IPv6 advanced ACL does not contain any rule. The vpn-instance <i>vpn-instance-name</i> option is not available on a 5500 SI switch. If an IPv6 advanced ACL is for QoS traffic classification or packet filtering: <ul style="list-style-type: none"> Do not specify the fragment, routing, or vpn-instance keyword, or specify neg for the <i>operator</i> argument. Do not specify the flow-label keyword if the ACL is for outbound QoS traffic classification or outbound packet filtering on a 5500 EI switch. The logging and counting keywords (even if specified) do not take effect for QoS traffic classification.
6. Add or edit a rule comment.	rule <i>rule-id</i> comment <i>text</i>	Optional. By default, no rule comments are configured.
7. Add or edit a rule range remark.	rule [<i>rule-id</i>] remark <i>text</i>	Optional. By default, no rule range remarks are configured.
8. Enable counting ACL rule matches performed in hardware.	hardware-count enable	Optional. Disabled by default. When the ACL is referenced by a QoS policy, this command does not take effect.

Configuring an Ethernet frame header ACL

Ethernet frame header ACLs, also called "Layer 2 ACLs," match packets based on Layer 2 protocol header fields, such as source MAC address, destination MAC address, 802.1p priority (VLAN priority), and link layer protocol type.

To configure an Ethernet frame header ACL:

Step	Command	Remarks
1. Enter system view.	<code>system-view</code>	N/A
2. Create an Ethernet frame header ACL and enter its view.	<code>acl number <i>acl-number</i></code> [<code>name <i>acl-name</i></code>] [<code>match-order { auto config }</code>]	By default, no ACL exists. Ethernet frame header ACLs are numbered in the range of 4000 to 4999. You can use the <code>acl name <i>acl-name</i></code> command to enter the view of a named Ethernet frame header ACL.
3. Configure a description for the Ethernet frame header ACL.	<code>description <i>text</i></code>	Optional. By default, an Ethernet frame header ACL has no ACL description.
4. Set the rule numbering step.	<code>step <i>step-value</i></code>	Optional. The default setting is 5.
5. Create or edit a rule.	<code>rule [<i>rule-id</i>] { deny permit } [cos <i>vlan-pri</i> counting dest-mac <i>dest-addr dest-mask</i> { isap <i>isap-type isap-type-mask</i> type <i>protocol-type protocol-type-mask</i> } source-mac <i>sour-addr source-mask</i> time-range <i>time-range-name</i>] *</code>	By default, an Ethernet frame header ACL does not contain any rule. If the ACL is for QoS traffic classification or packet filtering, to use the <code>isap</code> keyword, the <code>isap-type</code> argument must be AAAA, and the <code>isap-type-mask</code> argument must be FFFF. Otherwise, the ACL cannot be function normally.
6. Add or edit a rule comment.	<code>rule <i>rule-id</i> comment <i>text</i></code>	Optional. By default, no rule comments are configured.
7. Add or edit a rule range remark.	<code>rule [<i>rule-id</i>] remark <i>text</i></code>	Optional. By default, no rule range remarks are configured.
8. Enable counting ACL rule matches performed in hardware.	<code>hardware-count enable</code>	Optional. Disabled by default. When the ACL is referenced by a QoS policy, this command does not take effect.

Copying an ACL

You can create an ACL by copying an existing ACL (source ACL). The new ACL (destination ACL) has the same properties and content as the source ACL, but not the same ACL number and name.

To successfully copy an ACL, make sure that:

- The destination ACL number is from the same category as the source ACL number.
- The source ACL already exists but the destination ACL does not.

Copying an IPv4 ACL

Step	Command
1. Enter system view.	system-view
2. Copy an existing IPv4 ACL to create a new IPv4 ACL.	acl copy { <i>source-acl-number</i> name <i>source-acl-name</i> } to { <i>dest-acl-number</i> name <i>dest-acl-name</i> }

Copying an IPv6 ACL

Step	Command
1. Enter system view.	system-view
2. Copy an existing IPv6 ACL to generate a new one of the same category.	acl ipv6 copy { <i>source-acl6-number</i> name <i>source-acl6-name</i> } to { <i>dest-acl6-number</i> name <i>dest-acl6-name</i> }

Packet filtering with ACLs

You can use an ACL to filter incoming or outgoing IPv4 or IPv6 packets. You can apply one IPv4 ACL, one IPv6 ACL, and one Ethernet frame header ACL most to filter packets in the same direction of an interface.

With a basic or advanced ACL, you can log filtering events by specifying the **logging** keyword in the ACL rules and enabling the counting function. To enable counting for rule matches performed in hardware, configure the **hardware-count enable** command for the ACL or specify the **counting** keyword in the ACL rules.

You can set the packet filter to periodically send packet filtering logs to the information center as informational messages. The interval for generating and outputting packet filtering logs is configurable. The log information includes the number of matching packets and the ACL rules used in an interval. For more information about the information center, see *Network Management and Monitoring Configuration Guide*.

NOTE:

ACLs on VLAN interfaces filter only packets forwarded at Layer 3.

Applying an IPv4 or Ethernet frame header ACL for packet filtering

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Apply an IPv4 basic, IPv4 advanced, or Ethernet frame header ACL to the interface to filter packets.	packet-filter { <i>acl-number</i> name <i>acl-name</i> } { inbound outbound }	By default, no ACL is applied to any interface.
4. Exit to system view.	quit	N/A
5. Set the interval for generating and outputting IPv4 packet filtering logs.	acl logging frequency <i>frequency</i>	By default, the interval is 0. No IPv4 packet filtering logs are generated.

Applying an IPv6 ACL for packet filtering

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Apply an IPv6 basic or IPv6 advanced ACL to the interface to filter IPv6 packets.	packet-filter ipv6 { <i>acl6-number</i> name <i>acl6-name</i> } { inbound outbound }	By default, no IPv6 ACL is applied to the interface.
4. Exit to system view.	quit	N/A
5. Set the interval for generating and outputting IPv6 packet filtering logs.	acl ipv6 logging frequency <i>frequency</i>	The default interval is 0. No IPv6 packet filtering logs are generated.

Displaying and maintaining ACLs

Task	Command	Remarks
Display configuration and match statistics for one or all IPv4 ACLs.	display acl { <i>acl-number</i> all name <i>acl-name</i> } [<i>slot slot-number</i>] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display configuration and match statistics for one or all IPv6 ACLs.	display acl ipv6 { <i>acl6-number</i> all name <i>acl6-name</i> } [<i>slot slot-number</i>] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display the usage of ACL rules.	display acl resource [<i>slot slot-number</i>] [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view
Display the application status of packet filtering ACLs on interfaces.	display packet-filter { { all interface <i>interface-type interface-number</i> } [inbound outbound] interface vlan-interface <i>vlan-interface-number</i> [inbound outbound] } [<i>slot slot-number</i>] } [[{ begin exclude include } <i>regular-expression</i>]]	Available in any view

Task	Command	Remarks
Display the configuration and status of one or all time ranges.	display time-range { <i>time-range-name</i> all } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear statistics for one or all IPv4 ACLs.	reset acl counter { <i>acl-number</i> all name <i>acl-name</i> }	Available in user view
Clear statistics for one or all IPv6 basic and advanced ACLs.	reset acl ipv6 counter { <i>acl6-number</i> all name <i>acl6-name</i> }	Available in user view

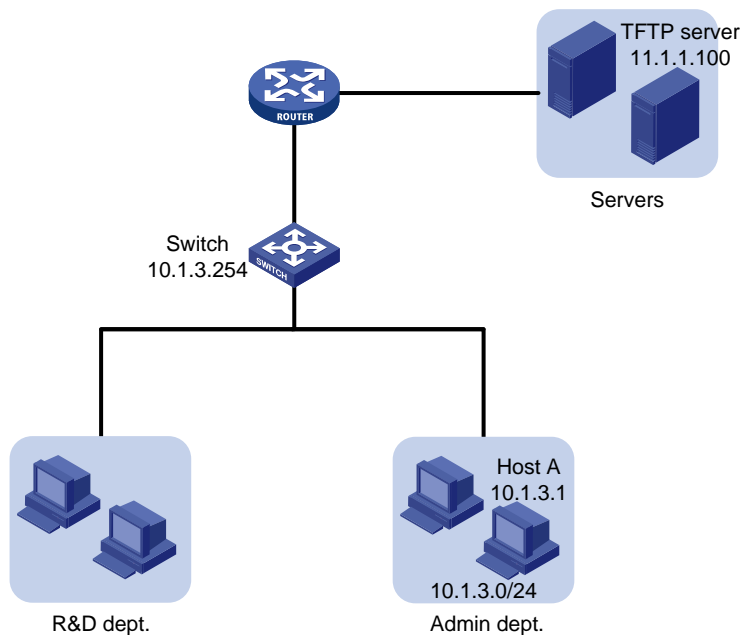
Configuration example of using ACL for device management

Network requirements

As shown in [Figure 1](#), configure ACLs so that:

- Host A can telnet to the switch only during the working time (8:30 to 18:00 of every working day).
- As a TFTP client, the switch can get files from only the server 11.1.1.100. This makes sure that the switch saves only authorized files.
- As an FTP server, the switch accepts the login requests from only the NMS.

Figure 1 Network diagram



Configuration procedure

1. Limit the telnet login requests.

Create a time range named **telnet** to cover 8:30 to 18:00 of every working day.

```
<Switch> system-view
```

```
[Switch] time-range telnet 8:30 to 18:00 working-day
# Create IPv4 basic ACL 2000, and configure a rule for the ACL to permit the packets sourced
from 10.1.3.1 during only the time specified by time range telnet.
[Switch] acl number 2000
[Switch-acl-basic-2000] rule permit source 10.1.3.1 0 time-range telnet
[Switch-acl-basic-2000] quit
# Apply ACL 2000 to the inbound traffic of all telnet user interfaces to limit the telnet login
requests.
[Switch] user-interface vty 0 4
[Switch-ui-vty0-4] acl 2000 inbound
```

2. Limit the access to the TFTP server.

Create IPv4 basic ACL 2001, and configure a rule for the ACL to permit only the packets sourced from 11.1.1.100.

```
[Switch] acl number 2001
[Switch-acl-basic-2001] rule permit source 11.1.1.100 0
[Switch-acl-basic-2001] quit
```

Use ACL 2001 to control the switch's access to a specific TFTP server.

```
[Switch] tftp-server acl 2001
```

3. Limit the FTP login requests.

Create IPv4 basic ACL 2002, and configure a rule for the ACL to permit only the packets sourced from 10.1.3.1.

```
[Switch] acl number 2002
[Switch-acl-basic-2001] rule permit source 10.1.3.1 0
[Switch-acl-basic-2001] quit
```

Enable the FTP server on the switch.

```
[Switch] ftp server enable
```

Use ACL 2001 to control FTP clients' access to the FTP server.

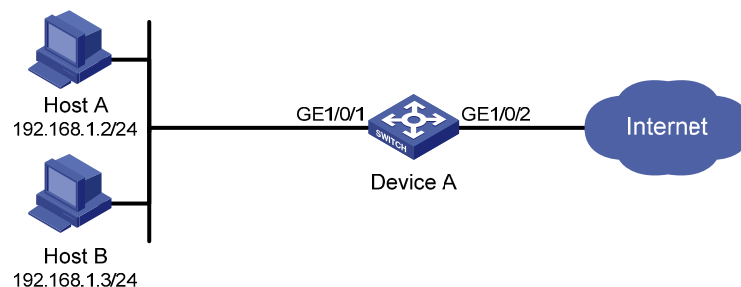
```
[Switch] ftp server acl 2002
```

IPv4 packet filtering configuration example

Network requirements

As shown in [Figure 2](#), apply an ACL to the inbound direction of interface GigabitEthernet 1/0/1 on Device A so that every day from 08:00 to 18:00 the interface allows only packets sourced from Host A to pass. Configure Device A to output IPv4 packet filtering logs to the console at 10-minute intervals.

Figure 2 Network diagram



Configuration procedure

Create a time range from 08:00 to 18:00 every day.

```
<DeviceA> system-view
[DeviceA] time-range study 8:00 to 18:00 daily
```

Create IPv4 ACL 2009, and configure two rules in the ACL. One rule permits packets sourced from Host A and the other denies packets sourced from any other host during the time range **study**. Enable logging for the permit rule.

```
[DeviceA] acl number 2009
[DeviceA-acl-basic-2009] rule permit source 192.168.1.2 0 time-range study logging
[DeviceA-acl-basic-2009] rule deny source any time-range study
[DeviceA-acl-basic-2009] quit
```

Enable the device to generate and output IPv4 packet filtering logs at 10-minute intervals.

```
[DeviceA] acl logging frequency 10
```

Configure the device to output informational log messages to the console.

```
[DeviceA] info-center source default channel 0 log level informational
```

Apply IPv4 ACL 2009 to filter incoming packets on GigabitEthernet 1/0/1.

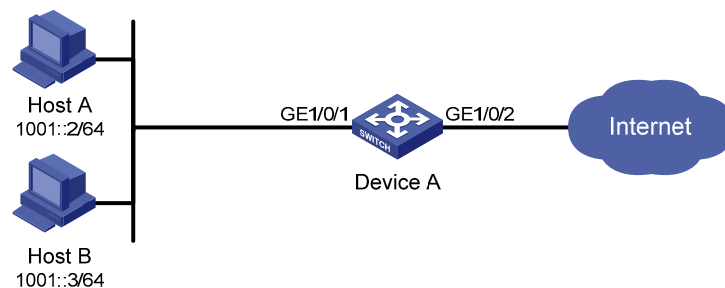
```
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] packet-filter 2009 inbound
[DeviceA-GigabitEthernet1/0/1] quit
```

IPv6 packet filtering configuration example

Network requirements

As shown in [Figure 3](#), apply an IPv6 ACL to the incoming traffic of GigabitEthernet 1/0/1 on Device A so that every day from 08:00 to 18:00 the interface allows only packets from Host A to pass through. Configure Device A to output IPv4 packet filtering logs to the console at 10-minute intervals.

Figure 3 Network diagram



Configuration procedure

Create a time range from 08:00 to 18:00 every day.

```
<DeviceA> system-view
[DeviceA] time-range study 8:0 to 18:0 daily
```

Create IPv6 ACL 2009, and configure two rules for the ACL. One permits packets sourced from Host A and the other denies packets sourced from any other host during the time range **study**. Enable logging for the permit rule.

```
[DeviceA] acl ipv6 number 2009
[DeviceA-acl6-basic-2009] rule permit source 1001::2 128 time-range study logging
[DeviceA-acl6-basic-2009] rule deny source any time-range study
[DeviceA-acl6-basic-2009] quit
```

Configure the device to collect and output IPv6 packet filtering logs at 10-minute intervals.

```
[DeviceA] acl ipv6 logging frequency 10
```

Configure the device to output informational log messages to the console.

```
[DeviceA] info-center source default channel 0 log level informational
```

Apply IPv6 ACL 2009 to filter incoming packets on GigabitEthernet 1/0/1.

```
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] packet-filter ipv6 2009 inbound
[DeviceA-GigabitEthernet1/0/1] quit
```

QoS overview

In data communications, Quality of Service (QoS) is a network's ability to provide differentiated service guarantees for diversified traffic in terms of bandwidth, delay, jitter, and drop rate.

Network resources are scarce. The contention for resources requires that QoS prioritize important traffic flows over trivial ones. For example, in the case of fixed bandwidth, if a traffic flow gets more bandwidth, the other traffic flows will get less bandwidth and may be affected. When making a QoS scheme, you must consider the characteristics of various applications to balance the interests of diversified users and to utilize network resources.

The following section describes some typical QoS service models and widely used, mature QoS techniques.

QoS service models

Best-effort service model

The best-effort model is a single-service model and also the simplest service model. In this service model, the network does its best to deliver packets, but does not guarantee delivery or control delay.

The best-effort service model is the default model in the Internet and applies to most network applications. It uses the first in first out (FIFO) queuing mechanism.

IntServ model

The integrated service (IntServ) model is a multiple-service model that can accommodate diverse QoS requirements. This service model provides the most granularly differentiated QoS by identifying and guaranteeing definite QoS for each data flow.

In the IntServ model, an application must request service from the network before it sends data. IntServ signals the service request with the Resource Reservation Protocol (RSVP). All nodes receiving the request reserve resources as requested and maintain state information for the application flow.

The IntServ model demands high storage and processing capabilities because it requires all nodes along the transmission path to maintain resource state information for each flow. This model is suitable for small-sized or edge networks, but not large-sized networks, for example, the core layer of the Internet, where billions of flows are present.

DiffServ model

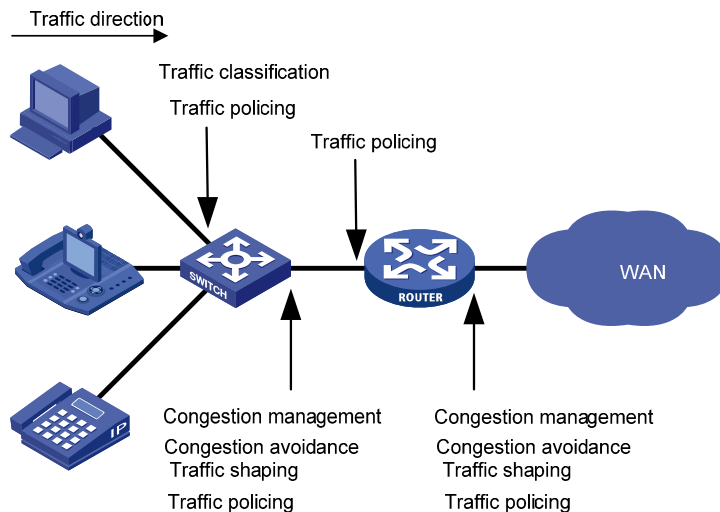
The differentiated service (DiffServ) model is a multiple-service model that can satisfy diverse QoS requirements. It is easy to implement and extend. DiffServ does not signal the network to reserve resources before sending data, as IntServ does.

All QoS techniques in this document are based on the DiffServ model.

QoS techniques

The QoS techniques include traffic classification, traffic policing, traffic shaping, line rate, congestion management, and congestion avoidance. They address problems that arise at different positions of a network.

Figure 4 Placement of the QoS techniques in a network



As shown in [Figure 4](#), traffic classification, traffic shaping, traffic policing, congestion management, and congestion avoidance mainly implement the following functions:

- **Traffic classification**—Uses certain match criteria to assign packets with the same characteristics to a class. Based on classes, you can provide differentiated services.
- **Traffic policing**—Policies flows entering or leaving a device, and imposes penalties on traffic flows that exceed the pre-set threshold to prevent aggressive use of network resources. You can apply traffic policing to both incoming and outgoing traffic of a port.
- **Traffic shaping**—Proactively adapts the output rate of traffic to the network resources available on the downstream device to eliminate packet drops. Traffic shaping usually applies to the outgoing traffic of a port.
- **Congestion management**—Provides a resource scheduling policy to determine the packet forwarding sequence when congestion occurs. Congestion management usually applies to the outgoing traffic of a port.
- **Congestion avoidance**—Monitors the network resource usage, and is usually applied to the outgoing traffic of a port. When congestion worsens, congestion avoidance reduces the queue length by dropping packets.

QoS configuration approaches

You can configure QoS in these approaches:

- [MQC approach](#)
- [Non-MQC approach](#)

Some features support both approaches, but some support only one.

MQC approach

In modular QoS configuration (MQC) approach, you configure QoS service parameters by using QoS policies (see "[Configuring a QoS policy](#)").

Non-MQC approach

In non-MQC approach, you configure QoS service parameters without using a QoS policy. For example, you can use the line rate feature to set a rate limit on an interface without using a QoS policy.

Configuring a QoS policy

Overview

A QoS policy is a set of class-behavior associations and defines the shaping, policing, or other QoS actions to take on different classes of traffic.

A class is a set of match criteria for identifying traffic and it uses the AND or OR operator:

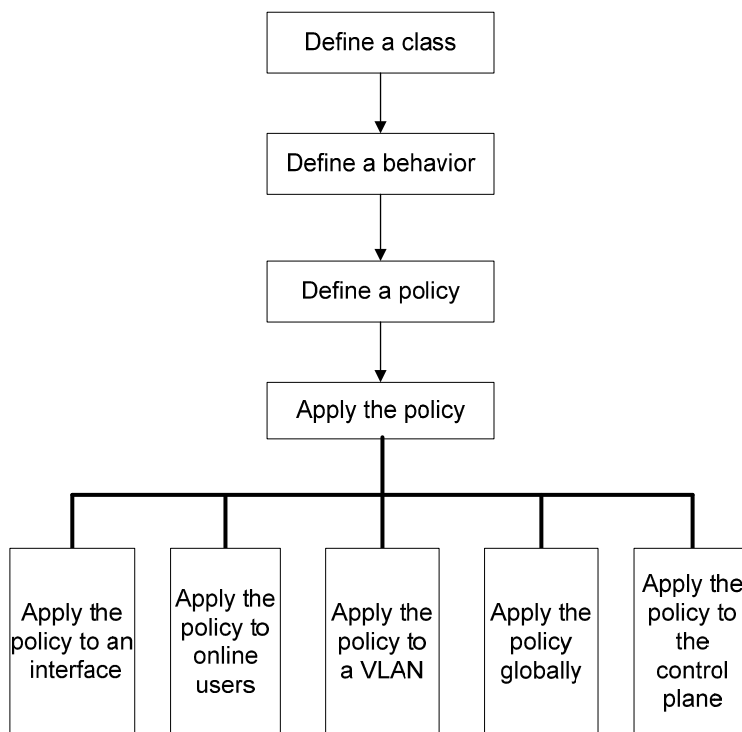
- **AND**—A packet must match all the criteria to match the class.
- **OR**—A packet matches the class if it matches any of the criteria in the class.

A traffic behavior defines a set of QoS actions to take on packets, such as priority marking and redirect.

By associating a traffic behavior with a class in a QoS policy, you apply the specific set of QoS actions to the class of traffic.

Figure 5 shows how to configure a QoS policy.

Figure 5 QoS policy configuration procedure



Defining a class

To define a class, specify its name and then configure the match criteria in class view.

Configuration restrictions and guidelines

- If a class that uses the AND operator has multiple **if-match acl**, **if-match acl ipv6**, **if-match customer-vlan-id** or **if-match service-vlan-id** clauses, a packet that matches any of the clauses matches the class.
- To successfully execute the traffic behavior associated with a traffic class that uses the AND operator, define only one **if-match** clause for any of the following match criteria and input only one value for any of the following *list* arguments. To create multiple **if-match** clauses for these match criteria or specify multiple values for the *list* arguments, specify the operator of the class as OR and use the **if-match** command multiple times.
 - **customer-dot1p** *8021p-list*
 - **destination-mac** *mac-address*
 - **dscp** *dscp-list*
 - **ip-precedence** *ip-precedence-list*
 - **service-dot1p** *8021p-list*
 - **source-mac** *mac-address*
 - **system-index** *index-value-list*

Configuration procedure

To define a class:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a class and enter class view.	traffic classifier <i>tcl-name</i> [operator { and or }]	<p>By default, the operator of a class is AND.</p> <p>The operator of a class can be AND or OR:</p> <ul style="list-style-type: none"> • AND—A packet is assigned to a class only when the packet matches all the criteria in the class. • OR—A packet is assigned to a class if it matches any of the criteria in the class.
3. Configure match criteria.	if-match <i>match-criteria</i>	N/A

match-criteria: Match criterion.

Table 2 The value range for the *match-criteria* argument

Option	Description
acl [ipv6] { <i>acl-number</i> name <i>acl-name</i> }	<p>Matches an ACL.</p> <p>The <i>acl-number</i> argument ranges from 2000 to 3999 for an IPv4 ACL, 2000 to 3999 for an IPv6 ACL, and 4000 to 4999 for an Ethernet frame header ACL.</p> <p>The <i>acl-name</i> argument is a case-insensitive string of 1 to 63 characters, which must start with an alphabetic letter from a to z (or A to Z), and to avoid confusion, cannot be all.</p>
any	Matches all packets.

Option	Description
dscp <i>dscp-list</i>	Matches DSCP values. The <i>dscp-list</i> argument is a list of up to eight DSCP values. A DSCP value can be a number from 0 to 63 or any keyword in Table 9 .
destination-mac <i>mac-address</i>	Matches a destination MAC address.
customer-dot1p <i>8021p-list</i>	Matches the 802.1p priority of the customer network. The <i>8021p-list</i> argument is a list of up to eight 802.1p priority values. An 802.1p priority ranges from 0 to 7.
service-dot1p <i>8021p-list</i>	Matches the 802.1p priority of the service provider network. The <i>8021p-list</i> argument is a list of up to eight 802.1p priority values. An 802.1p priority ranges from 0 to 7.
ip-precedence <i>ip-precedence-list</i>	Matches IP precedence. The <i>ip-precedence-list</i> argument is a list of up to eight IP precedence values. An IP precedence ranges from 0 to 7.
protocol <i>protocol-name</i>	Matches a protocol. The <i>protocol-name</i> argument can be IP or IPv6.
source-mac <i>mac-address</i>	Matches a source MAC address.
customer-vlan-id { <i>vlan-id-list</i> <i>vlan-id1 to vlan-id2</i> }	Matches the VLAN IDs of customer networks. The <i>vlan-id-list</i> argument is a list of up to eight VLAN IDs. The <i>vlan-id1 to vlan-id2</i> specifies a VLAN ID range, where the <i>vlan-id1</i> must be smaller than the <i>vlan-id2</i> . A VLAN ID ranges from 1 to 4094.
service-vlan-id { <i>vlan-id-list</i> <i>vlan-id1 to vlan-id2</i> }	Matches the VLAN IDs of ISP networks. The <i>vlan-id-list</i> is a list of up to eight VLAN IDs. The <i>vlan-id1 to vlan-id2</i> specifies a VLAN ID range, where the <i>vlan-id1</i> must be smaller than the <i>vlan-id2</i> . A VLAN ID ranges from 1 to 4094.
system-index <i>index-value-list</i>	Matches a pre-defined match criterion (system-index) for packets sent to the control plane. The <i>index-value-list</i> argument specifies a list of up to eight system indexes. The system index ranges from 1 to 128.

Defining a traffic behavior

A traffic behavior is a set of QoS actions (such as traffic filtering, shaping, policing, and priority marking) to take on a class of traffic. To define a traffic behavior, first create it and then configure QoS actions, such as priority marking and traffic redirecting, in traffic behavior view.

To define a traffic behavior:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a traffic behavior and enter traffic behavior view.	traffic behavior <i>behavior-name</i>	N/A
3. Configure actions in the traffic behavior.	See the subsequent chapters, depending on the purpose of the traffic behavior: traffic policing, traffic filtering, traffic redirecting, priority marking, traffic accounting, and so on.	

Defining a policy

You associate a behavior with a class in a QoS policy to perform the actions defined in the behavior for the class of packets.

Configuration restrictions and guidelines

- If an ACL is referenced by a QoS policy for defining traffic match criteria, packets matching the ACL are organized as a class and the behavior defined in the QoS policy applies to the class regardless of whether the action in the rule is **deny** or **permit**.
- In a QoS policy with multiple class-to-traffic-behavior associations, if the action of creating an outer VLAN tag, setting customer network VLAN ID, or setting service provider network VLAN ID is configured in a traffic behavior, do not configure any other action in this traffic behavior; otherwise, the QoS policy may not function as expected after it is applied. For more information about the action of setting customer network VLAN ID or service provider network VLAN ID, see *Layer 2—LAN Switching Configuration Guide*.

Configuration procedure

To associate a class with a behavior in a policy:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a policy and enter policy view.	qos policy <i>policy-name</i>	N/A
3. Associate a class with a behavior in the policy.	classifier <i>tcl-name</i> behavior <i>behavior-name</i> [mode dot1q-tag-manipulation]	Repeat this step to create more class-behavior associations.

The **dot1q-tag-manipulation** keyword is only for VLAN mapping purposes. For more information about VLAN mapping, see *Layer 2—LAN Switching Configuration Guide*.

Applying the QoS policy

You can apply a QoS policy to the following occasions:

- **An interface**—The policy takes effect on the traffic sent or received on the interface.
- **A user profile**—The policy takes effect on the traffic sent or received by the online users of the user profile.
- **A VLAN**—The policy takes effect on the traffic sent or received on all ports in the VLAN.
- **Globally**—The policy takes effect on the traffic sent or received on all ports.
- **Control plane**—The policy takes effect on the traffic received on the control plane.

The QoS policies applied to ports, to VLANs, and globally are in the descending priority order. If the system finds a matching QoS policy for the incoming/outgoing traffic, the system stops matching the traffic against QoS policies.

You can modify classes, behaviors, and class-behavior associations in a QoS policy applied to an interface, VLAN, or inactive user profile, or globally. If a class references an ACL for traffic classification, you can delete or modify the ACL (such as add rules to, delete rules from, and modify rules of the ACL).

If a QoS policy has been applied to an active user profile, you cannot modify classes, behaviors, and class-behavior associations of the QoS policy, or delete the QoS policy.

Applying the QoS policy to an interface

Both bridge mode (Layer 2) and route mode (Layer 3) Ethernet ports support QoS policies. The term *interface* in this section collectively refers to these types of ports. You can use the **port link-mode** command to set an Ethernet port to operate in bridge or route mode (see *Layer 2—LAN Switching Configuration Guide*).

The 5500 SI Switch Series does not support Layer 3 Ethernet ports.

A policy can be applied to multiple interfaces, but only one policy can be applied in one direction (inbound or outbound) of an interface.

The QoS policy applied to the outgoing traffic of a port does not regulate local packets, which are critical protocol packets sent by the device that hosts the interface for maintaining the normal operation of the device. The most common local packets include link maintenance packets, STP, LDP, and RSVP packets.

To apply the QoS policy to an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	<p>Use either command.</p> <p>Settings in interface view take effect on the current interface.</p> <p>Settings in port group view take effect on all ports in the port group.</p>
3. Apply the policy to the interface or port group.	qos apply policy <i>policy-name</i> { inbound outbound }	The outbound keyword is not available on the 5500 SI Switch Series.

Applying the QoS policy to online users

You can apply a QoS policy to multiple online users. In one direction of each online user, only one policy can be applied. To modify a QoS policy already applied in a certain direction, remove the QoS policy application first.

Configuration restrictions and guidelines

- The QoS policy applied to a user profile supports only the **remark**, **car**, and **filter** actions.
- Do not apply a null policy to a user profile. The user profile using a null policy cannot be activated.
- The authentication methods available for online users include 802.1X and Portal.

Configuration procedure

To apply the QoS policy to online users:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter user profile view.	user-profile <i>profile-name</i>	The configuration made in user profile view takes effect when the user profile is activated and the users of the user profile are online. For more information about user profiles, see <i>Security Configuration Guide</i> .
3. Apply the QoS policy.	qos apply policy <i>policy-name</i> { inbound outbound }	Use the inbound keyword to apply the QoS policy to the incoming traffic of the device (traffic sent by the online users). Use the outbound keyword to apply the QoS policy to the outgoing traffic (traffic received by the online users). The outbound keyword is not available on the 5500 SI Switch Series.
4. Return to system view.	quit	N/A
5. Activate the user profile.	user-profile <i>profile-name</i> enable	By default, a user profile is inactive.

Applying the QoS policy to a VLAN

You can apply a QoS policy to a VLAN to regulate traffic of the VLAN.

QoS policies cannot be applied to dynamic VLANs, such as VLANs created by GVRP.

To apply the QoS policy to a VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Apply the QoS policy to VLANs.	qos vlan-policy <i>policy-name</i> vlan <i>vlan-id-list</i> { inbound outbound }	The outbound keyword is not available on the 5500 SI Switch Series.

Applying the QoS policy globally

You can apply a QoS policy globally to the inbound or outbound direction of all ports.

To apply the QoS policy globally:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Apply the QoS policy globally.	qos apply policy <i>policy-name</i> global { inbound outbound }	The outbound keyword is not available on the 5500 SI Switch Series.

Applying the QoS policy to the control plane

A device provides the data plane and the control plane.

- The data plane has units responsible for receiving, transmitting, and switching (forwarding) packets, such as various dedicated forwarding chips. They deliver super processing speeds and throughput.
- The control plane has processing units running most routing and switching protocols and responsible for protocol packet resolution and calculation, such as CPUs. Compared with data plane units, the control plane units allow for great packet processing flexibility, but have lower throughput.

When the data plane receives packets that it cannot recognize or process, it transmits them to the control plane. If the transmission rate exceeds the processing capability of the control plane, which very likely occurs at times of DoS attacks, the control plane will be busy handling undesired packets and fail to handle legitimate packets correctly or timely. As a result, protocol performance is affected.

To address this problem, apply a QoS policy to the control plane to take QoS actions, such as traffic filtering or rate limiting, on inbound traffic. This action ensures that the control plane can receive, transmit, and process packets properly.

Configuration restrictions and guidelines

- By default, devices are configured with pre-defined control plane policies, which take effect on the control planes by default. A pre-defined control plane QoS policy uses the system-index to identify the type of packets sent to the control plane. You can reference system-indexes in **if-match** commands in class view for traffic classification and then re-configure traffic behaviors for these classes as required. You can use the **display qos policy control-plane pre-defined** command to display them.
- In a QoS policy for control planes, if a system index classifier is configured, the associated traffic behavior can contain only the **car** action or the combination of **car** and **accounting packet** actions. In addition, if the CAR action is configured, only its CIR setting can be applied.
- In the QoS policy for a control plane, if a system index classifier is not configured, the associated traffic behaviors also take effect on the data traffic of the device where the control plane resides.

Configuration procedure

To apply the QoS policy to the control plane:

Step	Command
1. Enter system view.	system-view
2. Enter control plane view.	control-plane slot <i>slot-number</i>
3. Apply the QoS policy to the control plane.	qos apply policy <i>policy-name</i> inbound

Displaying and maintaining QoS policies



IMPORTANT:

The **outbound** keyword is not available on the 5500 SI Switch Series.

Task	Command	Remarks
Display traffic class configuration.	display traffic classifier user-defined [<i>tcl-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display traffic behavior configuration.	display traffic behavior user-defined [<i>behavior-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display user-defined QoS policy configuration.	display qos policy user-defined [<i>policy-name</i> [classifier <i>tcl-name</i>]] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display QoS policy configuration on the specified or all interfaces.	display qos policy interface [<i>interface-type</i> <i>interface-number</i>] [inbound outbound] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display VLAN QoS policy configuration.	display qos vlan-policy { name <i>policy-name</i> vlan <i>vlan-id</i> } [slot <i>slot-number</i>] [inbound outbound] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about QoS policies applied globally.	display qos policy global [slot <i>slot-number</i>] [inbound outbound] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about QoS policies applied to a control plane.	display qos policy control-plane slot <i>slot-number</i> [inbound] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about pre-defined QoS policies applied to a control plane.	display qos policy control-plane pre-defined [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear VLAN QoS policy statistics.	reset qos vlan-policy [vlan <i>vlan-id</i>] [inbound outbound]	Available in user view
Clear the statistics for a QoS policy applied globally.	reset qos policy global [inbound outbound]	Available in user view
Clear the statistics for the QoS policy applied to a control plane.	reset qos policy control-plane slot <i>slot-number</i> [inbound]	Available in user view

Configuring priority mapping

Both bridge mode (Layer 2) and route mode (Layer 3) Ethernet ports support the priority mapping function. The term *interface* in this chapter collectively refers to these types of ports. You can use the **port link-mode** command to set an Ethernet port to operate in bridge or route mode (see *Layer 2—LAN Switching Configuration Guide*).

The 5500 SI Switch Series does not support Layer 3 Ethernet ports.

Overview

When a packet enters a device, depending on your configuration, the device assigns a set of QoS priority parameters to the packet based on either a certain priority field carried in the packet or the port priority of the incoming port. This process is called "priority mapping". During this process, the device can modify the priority of the packet depending on device status. The set of QoS priority parameters decides the scheduling priority and forwarding priority of the packet.

Priority mapping is implemented with priority mapping tables and involves priorities such as 802.1p priority, DSCP, IP precedence, local precedence, and drop precedence.

Types of priorities

Priorities fall into the following types: priorities carried in packets, and priorities locally assigned for scheduling only.

The packet-carried priorities include 802.1p priority, DSCP precedence, IP precedence, and so on. These priorities have global significance and affect the forwarding priority of packets across the network. For more information about these priorities, see "[Appendix B Packet precedences](#)."

The locally assigned priorities only have local significance. They are assigned by the device for scheduling only. These priorities include the local precedence and drop precedence, as follows:

- **Local precedence**—Local precedence is used for queuing. A local precedence value corresponds to an output queue. A packet with higher local precedence is assigned to a higher priority output queue to be preferentially scheduled.
- **Drop precedence**—Drop precedence is used for making packet drop decisions. Packets with the highest drop precedence are dropped preferentially.

Priority mapping tables

Priority mapping is implemented with priority mapping tables. By looking up a priority mapping table, the device decides which priority value to assign to a packet for subsequent packet processing. The switch provides the following priority mapping tables:

- **dot1p-dp**—802.1p-to-drop priority mapping table.
- **dot1p-lp**—802.1p-to-local priority mapping table.
- **dscp-dot1p**—DSCP-to-802.1p priority mapping table, which is applicable to only IP packets.
- **dscp-dp**—DSCP-to-drop priority mapping table, which is applicable to only IP packets.
- **dscp-dscp**—DSCP-to-DSCP priority mapping table, which is applicable to only IP packets.

The default priority mapping tables (see "[Appendix A Default priority mapping tables](#)") are available for priority mapping. In most cases, they are adequate for priority mapping. If a default priority mapping table cannot meet your requirements, you can modify the priority mapping table as required.

Priority trust mode on a port

The priority trust mode on a port decides which priority is used for priority mapping table lookup. Port priority was introduced to use for priority mapping in addition to priority fields carried in packets. The HP 5500 EI and 5500 SI Switch Series provides the following priority trust modes:

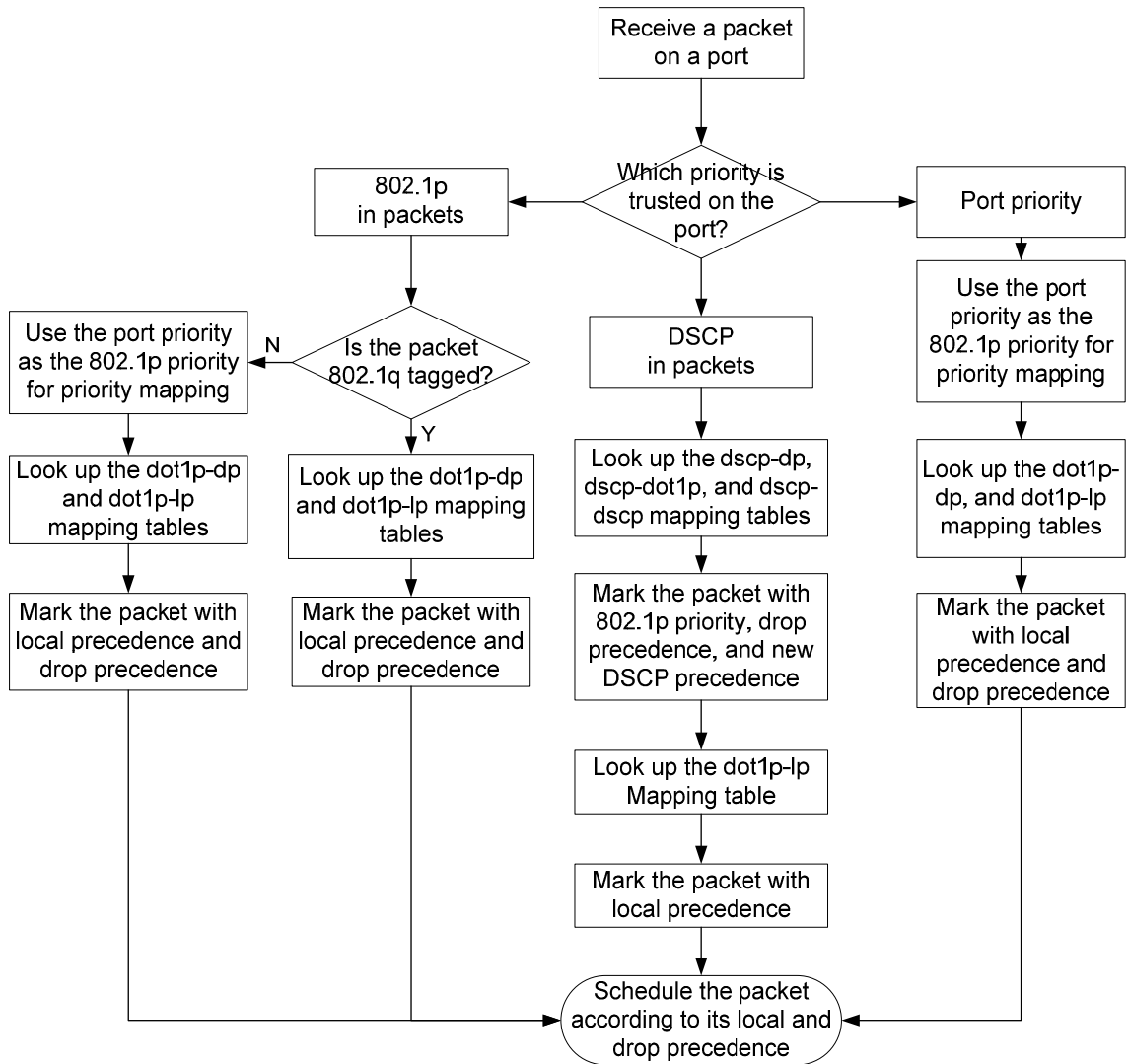
- Using the 802.1p priority carried in packets for priority mapping.
- Using the DSCP carried in packets for priority mapping.
- Using the port priority as the 802.1p priority for priority mapping. The port priority is user configurable.

The priority mapping procedure varies with the priority modes. For more information, see the subsequent section.

Priority mapping procedure

On receiving an Ethernet packet on a port, the switch marks the scheduling priorities (local precedence and drop precedence) for the Ethernet packet. This procedure is done according to the priority trust mode of the receiving port and the 802.1q tagging status of the packet, as shown in [Figure 6](#).

Figure 6 Priority mapping procedure for an Ethernet packet



The priority mapping procedure shown in Figure 6 applies in the absence of priority marking. If priority marking is configured, the switch performs priority marking before priority mapping. The switch then uses the re-marked packet-carried priority for priority mapping or directly uses the re-marked scheduling priority for traffic scheduling depending on your configuration. Neither priority trust mode configuration on the port nor port priority configuration takes effect.

Configuration guidelines

You can modify priority mappings by modifying priority mapping tables, priority trust mode on a port, and port priority.

HP recommends planning QoS throughout the network before making your QoS configuration.

Configuring a priority mapping table

To configure a priority mapping table:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter priority mapping table view.	qos map-table { dot1p-dp dot1p-lp dscp-dot1p dscp-dp dscp-dscp }	N/A
3. Configure the priority mapping table.	import <i>import-value-list</i> export <i>export-value</i>	Newly configured mappings overwrite the old ones.

Configuring a port to trust packet priority for priority mapping

When you configure the trusted packet priority type on an interface or port group, use the following priority trust modes:

- **dot1p**—Uses the 802.1p priority of received packets for mapping.
- **dscp**—Uses the DSCP precedence of received IP packets for mapping.
- **untrust**—Uses port priority as the 802.1p priority for priority mapping.

To configure the trusted packet priority type on an interface or port group:

Step	Command	Remarks
1. Enter system view	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> • Enter interface view: interface <i>interface-type</i> <i>interface-number</i> • Enter port group view: port-group manual <i>port-group-name</i> 	Use either command. Settings in interface view take effect on the current interface. Settings in port group view take effect on all ports in the port group.
3. Configure the trusted packet priority type for the interface.	<ul style="list-style-type: none"> • Trust the DSCP priority in packets: qos trust dscp • Trust the 802.1p priority in packets: qos trust dot1p • Trust the port priority: undo qos trust 	Use either command. By default, the device trusts the port priority.

Changing the port priority of an interface

To change the port priority of an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command. Settings in interface view take effect on the current interface. Settings in port group view take effect on all ports in the port group.
3. Set the port priority of the interface.	qos priority <i>priority-value</i>	The default port priority is 0.

Displaying priority mappings

Task	Command	Remarks
Display priority mapping table configuration.	display qos map-table [<i>dot1p-dp</i> <i>dot1p-lp</i> <i>dscp-dot1p</i> <i>dscp-dp</i> <i>dscp-dscp</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the trusted packet priority type on a port.	display qos trust interface [<i>interface-type</i> <i>interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

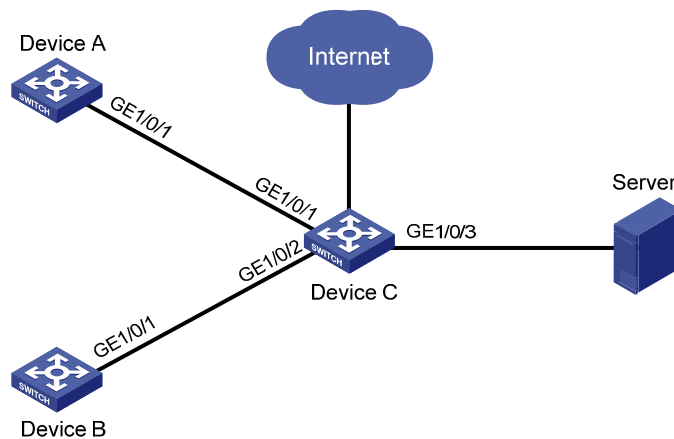
Priority trust mode configuration example

Network requirements

As shown in Figure 7, Device A is connected to GigabitEthernet 1/0/1 of Device C, Device B is connected to GigabitEthernet 1/0/2 of Device C, and the packets from Device A and Device B to Device C are not VLAN tagged.

Make configurations to have Device C preferentially process packets from Device A to Server when GigabitEthernet 1/0/3 of Device C is congested.

Figure 7 Network diagram



Configuration procedure

Assign port priority to GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2. Make sure that the priority of GigabitEthernet 1/0/1 is higher than that of GigabitEthernet 1/0/2, and no trusted packet priority type is configured on GigabitEthernet 1/0/1 or GigabitEthernet 1/0/2.

```
<DeviceC> system-view
[DeviceC] interface gigabitethernet 1/0/1
[DeviceC-GigabitEthernet1/0/1] qos priority 3
[DeviceC-GigabitEthernet1/0/1] quit
[DeviceC] interface gigabitethernet 1/0/2
[DeviceC-GigabitEthernet1/0/2] qos priority 1
[DeviceC-GigabitEthernet1/0/2] quit
```

Priority mapping table and priority marking configuration example

Network requirements

As shown in [Figure 8](#), the company's enterprise network interconnects all departments through Device. The network is described as follows:

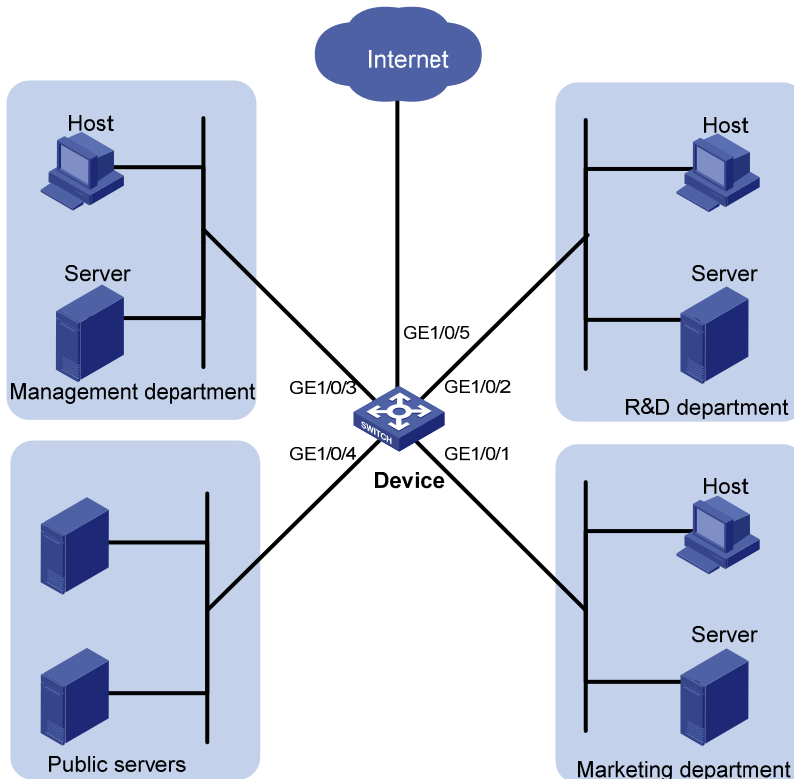
- The marketing department connects to GigabitEthernet 1/0/1 of Device, which sets the 802.1p priority of traffic from the marketing department to 3.
- The R&D department connects to GigabitEthernet 1/0/2 of Device, which sets the 802.1p priority of traffic from the R&D department to 4.
- The management department connects to GigabitEthernet 1/0/3 of Device, which sets the 802.1p priority of traffic from the management department to 5.

Configure port priority, 802.1p-to-local mapping table, and priority marking to implement the plan as described in [Table 3](#).

Table 3 Configuration plan

Traffic destination	Traffic priority order	Queuing plan		
		Traffic source	Output queue	Queue priority
Public servers	R&D department > management department > marketing department	R&D department	6	High
		Management department	4	Medium
		Marketing department	2	Low
Internet	Management department > marketing department > R&D department	R&D department	2	Low
		Management department	6	High
		Marketing department	4	Medium

Figure 8 Network diagram



Configuration procedure

1. Configure trusting port priority:

Set the port priority of GigabitEthernet 1/0/1 to 3.

```
<Device> system-view
[Device] interface gigabitethernet 1/0/1
[Device-GigabitEthernet1/0/1] qos priority 3
[Device-GigabitEthernet1/0/1] quit
```

Set the port priority of GigabitEthernet 1/0/2 to 4.

```
[Device] interface gigabitethernet 1/0/2
[Device-GigabitEthernet1/0/2] qos priority 4
[Device-GigabitEthernet1/0/2] quit
```

Set the port priority of GigabitEthernet 1/0/3 to 5.

```
[Device] interface gigabitethernet 1/0/3
[Device-GigabitEthernet1/0/3] qos priority 5
[Device-GigabitEthernet1/0/3] quit
```

2. Configure the priority mapping table:

Configure the 802.1p-to-local mapping table to map 802.1p priority values 3, 4, and 5 to local precedence values 2, 6, and 4. This guarantees the R&D department, management department, and marketing department decreased priorities to access the public server.

```
[Device] qos map-table dot1p-lp
[Device-maptbl-dot1p-lp] import 3 export 2
[Device-maptbl-dot1p-lp] import 4 export 6
```

```
[Device-maptbl-dot1p-1p] import 5 export 4
[Device-maptbl-dot1p-1p] quit
```

3. Configure priority marking:

Mark the HTTP traffic of the management department, marketing department, and R&D department to the Internet with 802.1p priorities 4, 5, and 3, respectively. Use the priority mapping table you have configured to map the 802.1p priorities to local precedence values 6, 4, and 2, respectively, for differentiated traffic treatment.

Create ACL 3000 to match HTTP traffic.

```
[Device] acl number 3000
[Device-acl-adv-3000] rule permit tcp destination-port eq 80
[Device-acl-adv-3000] quit
```

Create class **http** and reference ACL 3000 in the class.

```
[Device] traffic classifier http
[Device-classifier-http] if-match acl 3000
[Device-classifier-http] quit
```

Configure a priority marking policy for the management department, and apply the policy to the incoming traffic of GigabitEthernet 1/0/3.

```
[Device] traffic behavior admin
[Device-behavior-admin] remark dot1p 4
[Device-behavior-admin] quit
[Device] qos policy admin
[Device-qospolicy-admin] classifier http behavior admin
[Device-qospolicy-admin] quit
[Device] interface gigabitethernet 1/0/3
[Device-GigabitEthernet1/0/3] qos apply policy admin inbound
```

Configure a priority marking policy for the marketing department, and apply the policy to the incoming traffic of GigabitEthernet 1/0/1.

```
[Device] traffic behavior market
[Device-behavior-market] remark dot1p 5
[Device-behavior-market] quit
[Device] qos policy market
[Device-qospolicy-market] classifier http behavior market
[Device-qospolicy-market] quit
[Device] interface gigabitethernet 1/0/1
[Device-GigabitEthernet1/0/1] qos apply policy market inbound
```

Configure a priority marking policy for the R&D department, and apply the policy to the incoming traffic of GigabitEthernet 1/0/2.

```
[Device] traffic behavior rd
[Device-behavior-rd] remark dot1p 3
[Device-behavior-rd] quit
[Device] qos policy rd
[Device-qospolicy-rd] classifier http behavior rd
[Device-qospolicy-rd] quit
[Device] interface gigabitethernet 1/0/2
[Device-GigabitEthernet1/0/2] qos apply policy rd inbound
```

Configuring traffic policing, traffic shaping, and line rate

Both bridge mode (Layer 2) and route mode (Layer 3) Ethernet ports support the traffic shaping and line rate functions. The term *interface* in this chapter collectively refers to these types of ports. You can use the **port link-mode** command to set an Ethernet port to operate in bridge or route mode (see *Layer 2—LAN Switching Configuration Guide*).

The 5500 SI Switch Series does not support Layer 3 Ethernet ports.

Overview

Traffic policing, traffic shaping, and rate limit are QoS technologies that help assign network resources, such as assign bandwidth. They increase network performance and user satisfaction. For example, you can configure a flow to use only the resources committed to it in a certain time range. This avoids network congestion caused by burst traffic.

Traffic policing, generic traffic shaping (GTS), and line rate limit the traffic rate and resource usage according to traffic specifications. Once a particular flow exceeds its specifications, such as assigned bandwidth, the flow is shaped or policed to make sure that it is under the specifications. You can use token buckets for evaluating traffic specifications.

Traffic evaluation and token buckets

A token bucket is analogous to a container that holds a certain number of tokens. Each token represents a certain forwarding capacity. The system puts tokens into the bucket at a constant rate. When the token bucket is full, the extra tokens cause the token bucket to overflow.

Evaluating traffic with the token bucket

A token bucket mechanism evaluates traffic by looking at the number of tokens in the bucket. If the number of tokens in the bucket is enough for forwarding the packets, the traffic conforms to the specification, and is called "conforming traffic". Otherwise, the traffic does not conform to the specification, and is called "excess traffic".

A token bucket has the following configurable parameters:

- Mean rate at which tokens are put into the bucket, which is the permitted average rate of traffic. It is usually set to the committed information rate (CIR).
- Burst size or the capacity of the token bucket. It is the maximum traffic size permitted in each burst. It is usually set to the committed burst size (CBS). The set burst size must be greater than the maximum packet size.

Each arriving packet is evaluated. In each evaluation, if the number of tokens in the bucket is enough, the traffic conforms to the specification and the tokens for forwarding the packet are taken away; if the number of tokens in the bucket is not enough, the traffic is excessive.

Complicated evaluation

You can set two token buckets, bucket C and bucket E, to evaluate traffic in a more complicated environment and achieve more policing flexibility. For example, traffic policing uses the following parameters:

- **CIR**—Rate at which tokens are put into bucket C. It sets the average packet transmission or forwarding rate allowed by bucket C.
- **CBS**—Size of bucket C, which specifies the transient burst of traffic that bucket C can forward.
- **Peak information rate (PIR)**—Rate at which tokens are put into bucket E, which specifies the average packet transmission or forwarding rate allowed by bucket E.
- **Excess burst size (EBS)**—Size of bucket E, which specifies the transient burst of traffic that bucket E can forward.

CBS is implemented with bucket C, and EBS with bucket E. In each evaluation, packets are measured against the following bucket scenarios:

- If bucket C has enough tokens, packets are colored green.
- If bucket C does not have enough tokens but bucket E has enough tokens, packets are colored yellow.
- If neither bucket C nor bucket E has sufficient tokens, packets are colored red.

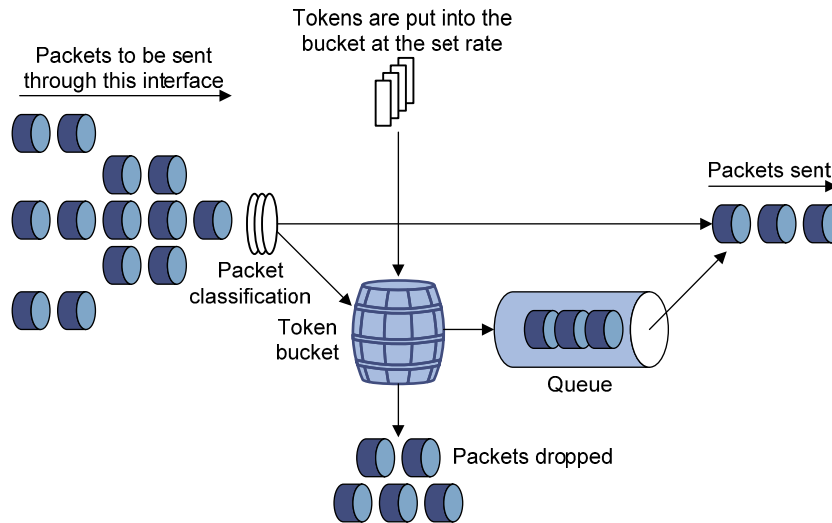
Traffic policing

ⓘ IMPORTANT:

The 5500 EI switch supports policing the inbound traffic and the outbound traffic, and the 5500 SI supports policing only the incoming traffic.

A typical application of traffic policing is to supervise the specification of certain traffic entering a network and limit it within a reasonable range, or to "discipline" the extra traffic to prevent aggressive use of network resources by a certain application. For example, you can limit bandwidth for HTTP packets to less than 50% of the total. If the traffic of a certain session exceeds the limit, traffic policing can drop the packets or reset the IP precedence of the packets. [Figure 9](#) shows an example of policing outbound traffic on an interface.

Figure 9 Traffic policing



Traffic policing is widely used in policing traffic entering the networks of internet service providers (ISPs). It can classify the policed traffic and take pre-defined policing actions on each packet depending on the evaluation result:

- Forwarding the packet if the evaluation result is "conforming"
- Dropping the packet if the evaluation result is "excess"
- Forwarding the packet with its precedence, which can be 802.1p priority (available only for green packets), DSCP, and local precedence, re-marked if the evaluation result is "conforming"

Traffic shaping

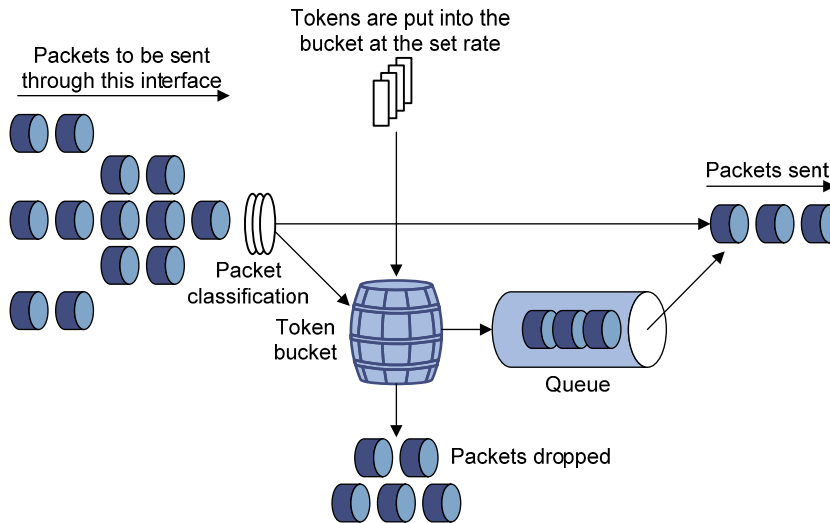
⚠ **IMPORTANT:**

Traffic shaping shapes the outbound traffic.

Traffic shaping limits the outbound traffic rate by buffering exceeding traffic. You can use traffic shaping to adapt the traffic output rate on a device to the input traffic rate of its connected device to avoid packet loss.

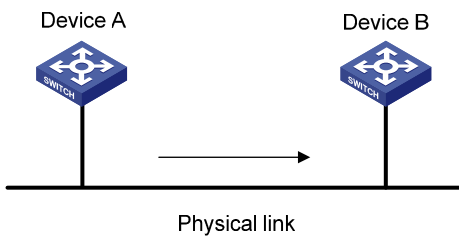
The difference between traffic policing and GTS is that packets to be dropped with traffic policing are retained in a buffer or queue with GTS, as shown in [Figure 10](#). When enough tokens are in the token bucket, the buffered packets are sent at an even rate. Traffic shaping can result in additional delay and traffic policing does not.

Figure 10 GTS



For example, in [Figure 11](#), Device B performs traffic policing on packets from Device A and drops packets exceeding the limit. To avoid packet loss, you can perform traffic shaping on the outgoing interface of Device A so packets exceeding the limit are cached in Device A. Once resources are released, traffic shaping takes out the cached packets and sends them out.

Figure 11 GTS application



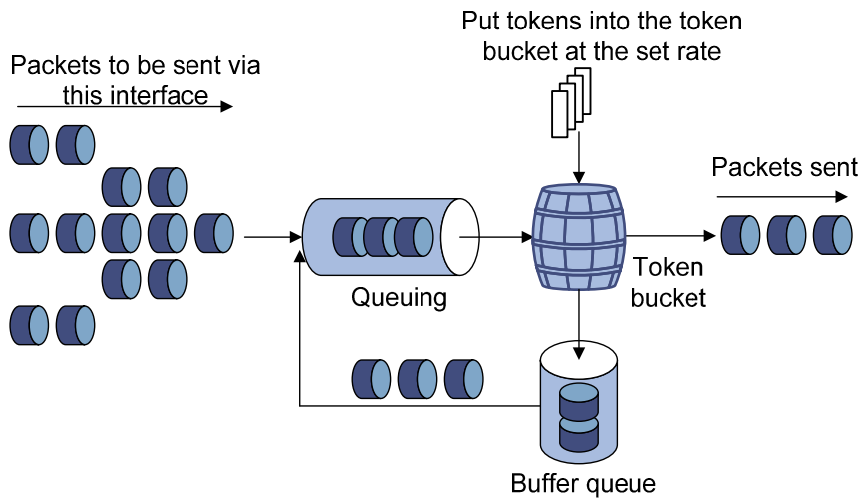
Line rate

Line rate supports rate-limiting the outbound traffic.

The line rate of a physical interface specifies the maximum rate for forwarding packets (including critical packets).

Line rate also uses token buckets for traffic control. With line rate configured on an interface, all packets to be sent through the interface are handled by the token bucket at line rate. If enough tokens are in the token bucket, packets can be forwarded. Otherwise, packets are put into QoS queues for congestion management. In this way, the traffic passing the physical interface is controlled.

Figure 12 Line rate implementation



The token bucket mechanism limits traffic rate when accommodating bursts. It allows bursty traffic to be transmitted if enough tokens are available. If tokens are scarce, packets cannot be transmitted until sufficient tokens are generated in the token bucket. It restricts the traffic rate to the rate for generating tokens.

Line rate can only limit traffic rate on a physical interface, and traffic policing can limit the rate of a flow on an interface. To limit the rate of all the packets on interfaces, using line rate is easier.

Configuring traffic policing

Configuration restrictions and guidelines

In a traffic behavior, do not configure traffic policing with any priority marking action (including local precedence, drop precedence, 802.1p priority, DSCP value, and IP precedence marking actions) in the same traffic behavior. Otherwise, you will fail to apply the QoS policy successfully.

Configuration procedure

To configure traffic policing:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a class and enter class view.	traffic classifier <i>tcl-name</i> [operator { and or }]	N/A
3. Configure match criteria.	if-match <i>match-criteria</i>	N/A
4. Return to system view.	quit	N/A
5. Create a behavior and enter behavior view.	traffic behavior <i>behavior-name</i>	N/A

Step	Command	Remarks
6. Configure a traffic policing action.	car cir <i>committed-information-rate</i> [cbs <i>committed-burst-size</i>] [ebs <i>excess-burst-size</i>] [pir <i>peak-information-rate</i>] [green action] [yellow action] [red action]	N/A
7. Return to system view.	quit	N/A
8. Create a policy and enter policy view.	qos policy <i>policy-name</i>	N/A
9. Associate the class with the traffic behavior in the QoS policy.	classifier <i>tcl-name</i> behavior <i>behavior-name</i>	N/A
10. Return to system view.	quit	N/A
11. Apply the QoS policy.	<ul style="list-style-type: none"> Applying the QoS policy to an interface Applying the QoS policy to online users Applying the QoS policy to a VLAN Applying the QoS policy globally Applying the QoS policy to the control plane 	Choose one application destination as needed.

Configuring GTS

The Switch Series supports queue-based GTS, which shapes traffic of a specific queue.

To configure GTS:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command. Settings in interface view take effect on the current interface. Settings in port group view take effect on all ports in the port group.
3. Configure GTS for a queue.	qos gts queue <i>queue-number</i> cir <i>committed-information-rate</i> [cbs <i>committed-burst-size</i>]	N/A

Configuring the line rate

The line rate of a physical interface specifies the maximum rate of outgoing packets.

To configure the line rate:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command. Settings in interface view take effect on the current interface. Settings in port group view take effect on all ports in the port group.
3. Configure the line rate for the interface or port group.	qos lr outbound cir <i>committed-information-rate</i> [cbs <i>committed-burst-size</i>]	N/A

Displaying and maintaining traffic policing, GTS, and line rate

On the 5500 EI and 5500 SI Switch Series, you can configure traffic policing in MQC approach. For more information about the displaying and maintaining commands, see "[Displaying and maintaining QoS policies.](#)"

Task	Command	Remarks
Display interface GTS configuration information.	display qos gts interface [<i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display interface line rate configuration information.	display qos lr interface [<i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Traffic policing configuration example

Network requirements

As shown in [Figure 13](#):

- GigabitEthernet 1/0/3 of Device A is connected to GigabitEthernet1/0/1 of Device B.
- Server, Host A, and Host B can access the Internet through Device A and Device B.

Perform traffic control on GigabitEthernet 1/0/1 of Device A for traffic received from Server and Host A, respectively, to satisfy the following requirements:

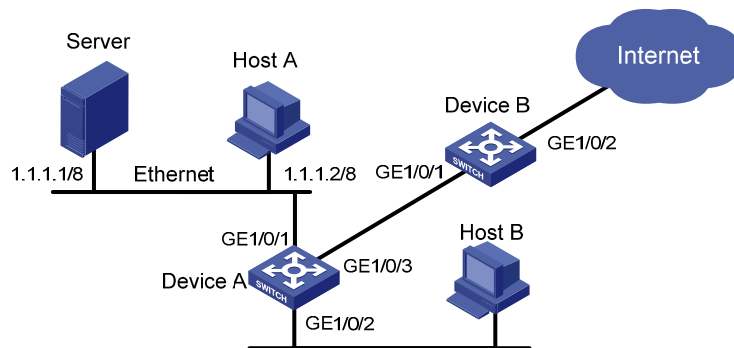
- Limit the rate of traffic from Server to 1024 kbps: transmit the conforming traffic normally, and mark the excess traffic with DSCP value 0 and then transmit the traffic.
- Limit the rate of traffic from Host A to 256 kbps: transmit the conforming traffic normally, and drop the excess traffic.

Perform traffic control on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 of Device B to satisfy the following requirements:

- Limit the total incoming traffic rate of GigabitEthernet 1/0/1 to 2048 kbps, and drop the excess traffic.

- Limit the outgoing HTTP traffic (traffic accessing the Internet) rate of GigabitEthernet 1/0/2 to 1024 kbps, and drop the excess traffic.

Figure 13 Network diagram



Configuration procedures

1. Configure Device A:

Configure ACL 2001 and ACL 2002 to match traffic from Server and Host A, respectively.

```
<DeviceA> system-view
[DeviceA] acl number 2001
[DeviceA-acl-basic-2001] rule permit source 1.1.1.1 0
[DeviceA-acl-basic-2001] quit
[DeviceA] acl number 2002
[DeviceA-acl-basic-2002] rule permit source 1.1.1.2 0
[DeviceA-acl-basic-2002] quit
```

Create a class named **server**, and use ACL 2001 as the match criterion. Create a class named **host**, and use ACL 2002 as the match criterion.

```
[DeviceA] traffic classifier server
[DeviceA-classifier-server] if-match acl 2001
[DeviceA-classifier-server] quit
[DeviceA] traffic classifier host
[DeviceA-classifier-host] if-match acl 2002
[DeviceA-classifier-host] quit
```

Create a behavior named **server**, and configure the CAR action for the behavior as follows: set the CIR to 1024 kbps, and mark the excess packets (red packets) with DSCP value 0 and transmit them.

```
[DeviceA] traffic behavior server
[DeviceA-behavior-server] car cir 1024 red remark-dscp-pass 0
[DeviceA-behavior-server] quit
```

Create a behavior named **host**, and configure the CAR action for the behavior as follows: set the CIR to 256 kbps.

```
[DeviceA] traffic behavior host
[DeviceA-behavior-host] car cir 256
[DeviceA-behavior-host] quit
```

Create a QoS policy named **car**, and associate class **server** with behavior **server** and class **host** with behavior **host**.

```
[DeviceA] qos policy car
```

```
[DeviceA-qospolicy-car] classifier server behavior server
[DeviceA-qospolicy-car] classifier host behavior host
[DeviceA-qospolicy-car] quit
# Apply QoS policy car to the incoming traffic of port GigabitEthernet 1/0/1.
[DeviceA] interface GigabitEthernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] qos apply policy car inbound
```

2. Configure Device B:

```
# Configure advanced ACL 3001 to match HTTP traffic.
<DeviceB> system-view
[DeviceB] acl number 3001
[DeviceB-acl-adv-3001] rule permit tcp destination-port eq 80
[DeviceB-acl-adv-3001] quit

# Create a class named http, and use ACL 3001 as the match criterion.
[DeviceB] traffic classifier http
[DeviceB-classifier-http] if-match acl 3001
[DeviceB-classifier-http] quit

# Create a class named class, and configure the class to match all packets.
[DeviceB] traffic classifier class
[DeviceB-classifier-class] if-match any
[DeviceB-classifier-class] quit

# Create a behavior named car_inbound, and configure the CAR action for the behavior as follows: set the CIR to 2048 kbps.
[DeviceB] traffic behavior car_inbound
[DeviceB-behavior-car_inbound] car cir 2048
[DeviceB-behavior-car_inbound] quit

# Create a behavior named car_outbound, and configure a CAR action for the behavior as follows: set the CIR to 1024 kbps.
[DeviceB] traffic behavior car_outbound
[DeviceB-behavior-car_outbound] car cir 1024
[DeviceB-behavior-car_outbound] quit

# Create a QoS policy named car_inbound, and associate class class with traffic behavior car_inbound in the QoS policy.
[DeviceB] qos policy car_inbound
[DeviceB-qospolicy-car_inbound] classifier class behavior car_inbound
[DeviceB-qospolicy-car_inbound] quit

# Create a QoS policy named car_outbound, and associate class http with traffic behavior car_outbound in the QoS policy.
[DeviceB] qos policy car_outbound
[DeviceB-qospolicy-car_outbound] classifier http behavior car_outbound
[DeviceB-qospolicy-car_outbound] quit

# Apply QoS policy car_inbound to the incoming traffic of port GigabitEthernet 1/0/1.
[DeviceB] interface GigabitEthernet 1/0/1
[DeviceB-GigabitEthernet1/0/1] qos apply policy car_inbound inbound

# Apply QoS policy car_outbound to the outgoing traffic of port GigabitEthernet 1/0/2.
[DeviceB] interface GigabitEthernet 1/0/2
[DeviceB-GigabitEthernet1/0/2] qos apply policy car_outbound outbound
```


Configuring congestion management

Both bridge mode (Layer 2) and route mode (Layer 3) Ethernet ports support the congestion management function. The term *interface* in this chapter collectively refers to these types of ports. You can use the **port link-mode** command to set an Ethernet port to operate in bridge or route mode (see *Layer 2—LAN Switching Configuration Guide*).

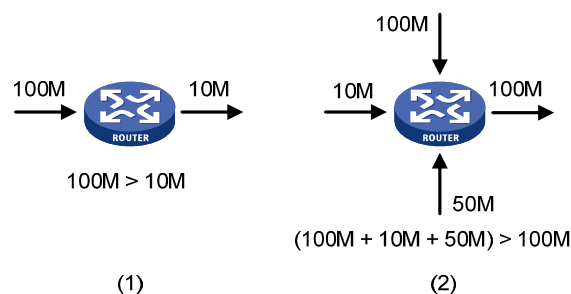
The 5500 SI Switch Series does not support Layer 3 Ethernet ports.

Overview

Network congestion degrades service quality on a traditional network. Congestion is a situation where the forwarding rate decreases due to insufficient resources, resulting in extra delay.

Congestion is more likely to occur in complex packet switching circumstances. Figure 14 shows two common cases:

Figure 14 Traffic congestion causes



Congestion can bring the following negative results:

- Increased delay and jitter during packet transmission
- Decreased network throughput and resource use efficiency
- Network resource (memory, in particular) exhaustion and system breakdown

Congestion is unavoidable in switched networks and multi-user application environments. To improve the service performance of your network, you must take proper measures to address the congestion issues.

The key to congestion management is defining a dispatching policy for resources to decide the order of forwarding packets when congestion occurs.

Congestion management techniques

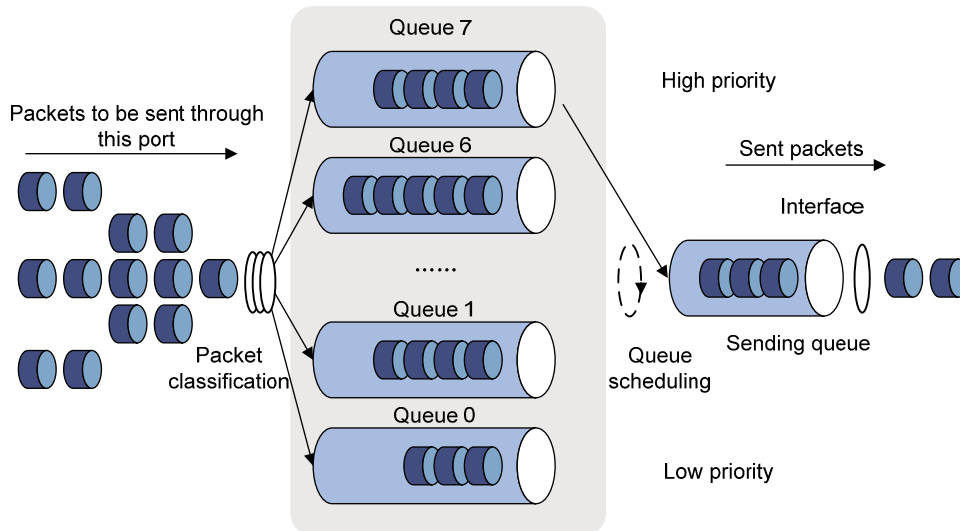
Congestion management uses queuing and scheduling algorithms to classify and sort traffic leaving a port. Each queuing algorithm addresses a particular network traffic problem, and has a different impact on bandwidth resource assignment, delay, and jitter.

Queue scheduling processes packets by their priorities, preferentially forwarding high-priority packets. The following section describes Strict Priority (SP) queuing, Weighted Fair Queuing (WFQ), Weighted Round Robin (WRR) queuing, SP+WRR queuing, and SP+WFQ queuing.

SP queuing

SP queuing is designed for mission-critical applications that require preferential service to reduce the response delay when congestion occurs.

Figure 15 SP queuing



In Figure 15, SP queuing classifies eight queues on a port into eight classes, numbered 7 to 0 in descending priority order.

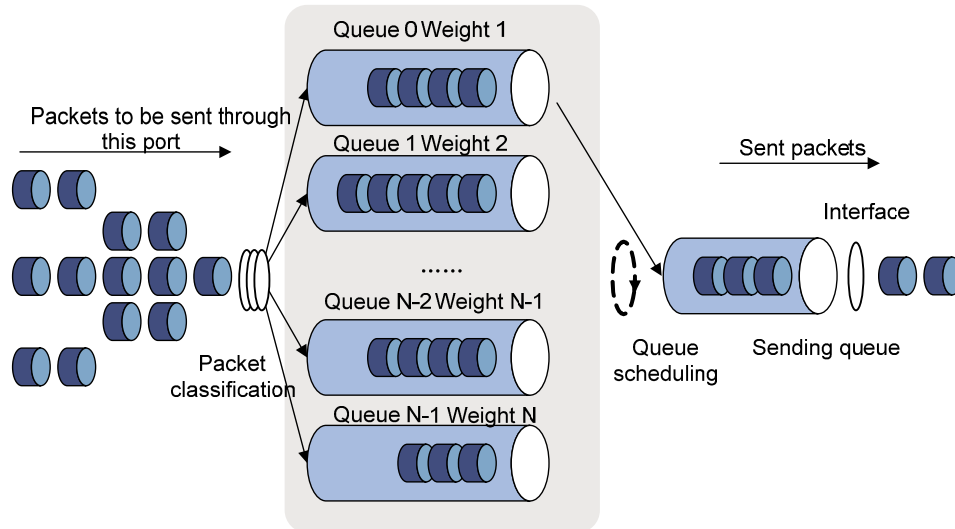
SP queuing schedules the eight queues in the descending order of priority. SP queuing sends packets in the queue with the highest priority first. When the queue with the highest priority is empty, it sends packets in the queue with the second highest priority, and so on. You can assign mission-critical packets to the high priority queue to make sure that they are always served first, and assign common service packets to the low priority queues and transmitted when the high priority queues are empty.

The disadvantage of SP queuing is that packets in the lower priority queues cannot be transmitted if packets exist in the higher priority queues. This may cause lower priority traffic to starve to death.

WRR queuing

WRR queuing schedules all the queues in turn to ensure every queue is served for a certain time, as shown in Figure 16.

Figure 16 WRR queuing



Assume a port provides eight output queues. WRR assigns each queue a weight value (represented by $w_7, w_6, w_5, w_4, w_3, w_2, w_1$, or w_0) to decide the proportion of resources assigned to the queue.

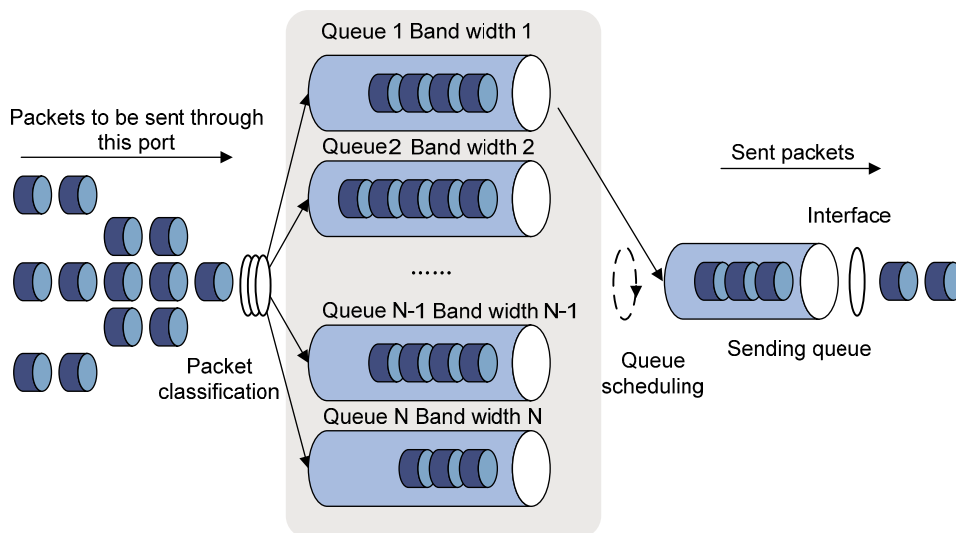
- The 5500 SI switch supports byte-count weight, which determines the weight by the number of bytes scheduled in a cycle.
- The 5500 EI switch supports byte-count weight (which determines the weight by the number of bytes scheduled in a cycle) or packet-based weight (which determines the weight by the number of packets scheduled in a cycle).

Take the byte-count weight as an example. On a 1000 Mbps port, you can configure the weight values of WRR queuing to 5, 5, 3, 3, 1, 1, 1, and 1 (corresponding to $w_7, w_6, w_5, w_4, w_3, w_2, w_1$, and w_0 , respectively). In this way, the queue with the lowest priority can get a minimum of 50 Mbps of bandwidth. WRR avoids the disadvantage of SP queuing, where packets in low-priority queues can fail to be served for a long time.

Another advantage of WRR queuing is that when the queues are scheduled in turn, the service time for each queue is not fixed. If a queue is empty, the next queue will be scheduled immediately. This improves bandwidth resource use efficiency.

WFQ queuing

Figure 17 WFQ queuing



WFQ is similar to WRR. You can use WFQ as an alternative to WRR.

Compared with WRR, WFQ can work with the minimum guaranteed bandwidth as follows:

- By setting the minimum guaranteed bandwidth, you can make sure that each WFQ queue is assured of certain bandwidth.
- The assignable bandwidth is allocated based on the priority of each queue (assignable bandwidth = total bandwidth – the sum of minimum guaranteed bandwidth of each queue).

For example, assume the total bandwidth of a port is 10 Mbps, and the port has five flows, with the precedence being 0, 1, 2, 3, and 4 and the minimum guaranteed bandwidth being 128 kbps, 128 kbps, 128 kbps, 64 kbps, and 64 kbps, respectively.

- The assignable bandwidth = 10 Mbps – (128 kbps + 128 kbps + 128 kbps + 64 kbps + and 64 kbps) = 9.5 Mbps.
- The total assignable bandwidth quota is the sum of all the (precedence value + 1)s, 1 + 2 + 3 + 4 + 5 = 15.
- The bandwidth percentage assigned to each flow is (precedence value of the flow + 1)/total assignable bandwidth quota. The bandwidth percentages for the flows are 1/15, 2/15, 3/15, 4/15, and 5/15, respectively.
- The bandwidth assigned to a queue = the minimum guaranteed bandwidth + the bandwidth allocated to the queue from the assignable bandwidth.

SP+WRR queuing

You can assign some queues on a port to the SP scheduling group and the others to the WRR scheduling group (group 1) to implement SP + WRR queue scheduling. The switch schedules packets in the SP scheduling group preferentially, and when the SP scheduling group is empty, schedules the packets in the WRR scheduling group. Queues in the SP scheduling group are scheduled with the SP queue scheduling algorithm. Queues in the WRR scheduling group are scheduled with WRR.

SP+WFQ queuing

SP+WFQ queuing is similar to SP+WRR queuing. You can assign some queues on a port to the SP scheduling group and the others to the WFQ scheduling group to implement SP + WFQ queue scheduling. The switch schedules packets of queues in the WFQ group based on their minimum guaranteed bandwidth settings, then uses SP queuing to schedule the queues in the SP scheduling group, and at last uses WFQ to schedule the queues in the WFQ scheduling group in a round robin fashion according to their weights

Configuring SP queuing

Configuration procedure

To configure SP queuing:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none">Enter interface view: interface <i>interface-type</i> <i>interface-number</i>Enter port group view: port-group manual <i>port-group-name</i>	Use either command. Settings in interface view take effect on the current interface. Settings in port group view take effect on all ports in the port group.
3. Configure SP queuing.	qos sp	The default queuing algorithm on an interface is WRR queuing.
4. Display SP queuing configuration.	display qos sp interface [<i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Optional. Available in any view

Configuration example

Network requirements

Configure GigabitEthernet 1/0/1 to use SP queuing.

Configuration procedure

```
# Enter system view
<Sysname> system-view

# Configure GigabitEthernet1/0/1 to use SP queuing.
[Sysname] interface gigabitethernet 1/0/1
[Sysname-GigabitEthernet1/0/1] qos sp
```

Configuring WRR queuing

Configuration procedure

To configure WRR queuing:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none">Enter interface view: interface <i>interface-type</i> <i>interface-number</i>Enter port group view: port-group manual <i>port-group-name</i>	Use either command. Settings in interface view take effect on the current interface. Settings in port group view take effect on all ports in the port group.
3. Enable byte-count or packet-based WRR queuing.	qos wrr [byte-count weight]	Optional. The default queuing algorithm on an interface is WRR. Only the 5500 EI switch supports the byte-count and weight keywords.
4. Configure the scheduling weight for a queue.	<ul style="list-style-type: none">For a byte-count WRR queue: qos wrr <i>queue-id</i> group <i>group-id</i> byte-count <i>schedule-value</i>For a packet-based WRR queue: qos wrr <i>queue-id</i> group <i>group-id</i> weight <i>schedule-value</i>	Select an approach according to the WRR queuing type. The 5500 SI switch supports only packet-based WRR queue configuration. By default, packet-based WRR is used, and the weights of queues 0 through 7 are 1, 2, 3, 4, 5, 9, 13, and 15.
5. Display WRR queuing configuration information on interfaces.	display qos wrr interface [<i>interface-type</i> <i>interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Optional. Available in any view

NOTE:

To guarantee successful WRR configuration, make sure that the scheduling weight type (byte-count or packet-based) is the same as the WRR queuing type (byte-count or packet-based) when you configure the scheduling weight for a WRR queue.

Configuration example

WRR queuing configuration example on a 5500 EI switch

- Network requirements
 - Configure byte-count WRR on GigabitEthernet 1/0/1.
 - Assign all queues to the WRR group, with the weights of 1, 2, 4, 6, 8, 10, 12, and 14.

2. Configuration procedures

Enter system view.

```
<Sysname> system-view
```

Configure WRR queuing on port GigabitEthernet 1/0/1.

```
[Sysname] interface GigabitEthernet 1/0/1
[Sysname-GigabitEthernet1/0/1] qos wrr byte-count
[Sysname-GigabitEthernet1/0/1] qos wrr 0 group 1 byte-count 1
[Sysname-GigabitEthernet1/0/1] qos wrr 1 group 1 byte-count 2
[Sysname-GigabitEthernet1/0/1] qos wrr 2 group 1 byte-count 4
[Sysname-GigabitEthernet1/0/1] qos wrr 3 group 1 byte-count 6
[Sysname-GigabitEthernet1/0/1] qos wrr 4 group 1 byte-count 8
[Sysname-GigabitEthernet1/0/1] qos wrr 5 group 1 byte-count 10
[Sysname-GigabitEthernet1/0/1] qos wrr 6 group 1 byte-count 12
[Sysname-GigabitEthernet1/0/1] qos wrr 7 group 1 byte-count 14
```

WRR queuing configuration example on a 5500 SI switch

1. Network requirements

- Configure WRR queuing on port GigabitEthernet 1/0/1.
- Assign all queues to the WRR group, with the weights of 1, 2, 4, 6, 8, 10, 12, and 14.

2. Configuration procedures

Enter system view.

```
<Sysname> system-view
```

Configure WRR queuing on port GigabitEthernet 1/0/1.

```
[Sysname] interface GigabitEthernet 1/0/1
[Sysname-GigabitEthernet1/0/1] qos wrr
[Sysname-GigabitEthernet1/0/1] qos wrr 0 group 1 weight 1
[Sysname-GigabitEthernet1/0/1] qos wrr 1 group 1 weight 2
[Sysname-GigabitEthernet1/0/1] qos wrr 2 group 1 weight 4
[Sysname-GigabitEthernet1/0/1] qos wrr 3 group 1 weight 6
[Sysname-GigabitEthernet1/0/1] qos wrr 4 group 1 weight 8
[Sysname-GigabitEthernet1/0/1] qos wrr 5 group 1 weight 10
[Sysname-GigabitEthernet1/0/1] qos wrr 6 group 1 weight 12
[Sysname-GigabitEthernet1/0/1] qos wrr 7 group 1 weight 14
```

Configuring WFQ queuing

Configuration procedure

To configure WFQ queuing:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command. Settings in interface view take effect on the current interface. Settings in port group view take effect on all ports in the port group.
3. Enable byte-count or packet-based WFQ queuing.	qos wfq [byte-count weight]	The default queuing algorithm on an interface is WRR.
4. Configure the scheduling weight for a queue.	<ul style="list-style-type: none"> For a byte-count WFQ queue: qos wfq <i>queue-id</i> group <i>group-id</i> byte-count <i>schedule-value</i> For a packet-based WFQ queue: qos wfq <i>queue-id</i> group <i>group-id</i> weight <i>schedule-value</i> 	Select a command according to the WFQ type (byte-count or packet-based) you have enabled. If you have enabled WFQ on the port, byte-count WRR applies by default, and the default scheduling weight is 1 for each queue.
5. Configure the minimum guaranteed bandwidth for a WFQ queue.	qos bandwidth queue <i>queue-id</i> min <i>bandwidth-value</i>	Optional. 64 kbps by default for each queue.
6. Display WFQ queuing configuration.	display qos wfq interface [<i>interface-type</i> <i>interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Optional. Available in any view

NOTE:

To guarantee successful WFQ configuration, make sure that the scheduling weight type (byte-count or packet-based) is the same as the WFQ queuing type (byte-count or packet-based) when you configure the scheduling weight for a WFQ queue.

Configuration example

Network requirements

Configure WFQ queues on an interface and assign the scheduling weight 2, 5, 10, and 10 to queue 1, queue 3, queue 4, queue 5, and queue 6, respectively.

Configuration procedure

```
# Enter system view.
<Sysname> system-view

# Configure WFQ queues on GigabitEthernet 1/0/1.
[Sysname] interface gigabitethernet 1/0/1
[Sysname-GigabitEthernet1/0/1] qos wfq
[Sysname-GigabitEthernet1/0/1] qos wfq 1 weight 2
[Sysname-GigabitEthernet1/0/1] qos wfq 3 weight 5
[Sysname-GigabitEthernet1/0/1] qos wfq 4 weight 10
[Sysname-GigabitEthernet1/0/1] qos wfq 5 weight 10
[Sysname-GigabitEthernet1/0/1] qos wfq 6 weight 10
```


Configuring SP+WRR queuing

Configuration procedure

To configure SP + WRR queuing:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none">Enter interface view: interface <i>interface-type</i> <i>interface-number</i>Enter port group view: port-group manual <i>port-group-name</i>	Use either command. Settings in interface view take effect on the current interface. Settings in port group view take effect on all ports in the port group.
3. Enable byte-count or packet-based WRR queuing.	qos wrr [byte-count weight]	Optional. By default, all ports use WRR queuing. Only the 5500 EI switch supports the byte-count and weight keywords.
4. Configure SP queue scheduling.	qos wrr queue-id group sp	By default, all the queues of a WRR-enabled port use the WRR queue scheduling algorithm.
5. Assign a queue to a WRR group and configure the scheduling weight for the queue.	qos wrr queue-id group <i>group-id</i> { weight byte-count } <i>schedule-value</i>	By default, on a WRR-enabled port, packet-based WRR is enabled, and the weights of queues 0 through 7 are 1, 2, 3, 4, 5, 9, 13, and 15. The 5500 SI switch supports only the weight keyword.

NOTE:

To guarantee successful WRR configuration, make sure that the scheduling weight type (byte-count or packet-based) is the same as the WRR queuing type (byte-count or packet-based) when you configure the scheduling weight for a WRR queue.

Configuration example

Network requirements

- Configure SP+WRR queue scheduling algorithm on GigabitEthernet 1/0/1, and use packet-based WRR.
- Configure queue 0, queue 1, queue 2, and queue 3 on GigabitEthernet 1/0/1 to be in SP queue scheduling group.
- Configure queue 4, queue 5, queue 6, and queue 7 on GigabitEthernet 1/0/1 to use WRR queuing, with the weight 2, 4, 6, and 8, respectively.

Configuration procedure

```
# Enter system view.
<Sysname> system-view

# Enable the SP+WRR queue scheduling algorithm on GigabitEthernet1/0/1.
[Sysname] interface GigabitEthernet 1/0/1
[Sysname-GigabitEthernet1/0/1] qos wrr weight
[Sysname-GigabitEthernet1/0/1] qos wrr 0 group sp
[Sysname-GigabitEthernet1/0/1] qos wrr 1 group sp
[Sysname-GigabitEthernet1/0/1] qos wrr 2 group sp
[Sysname-GigabitEthernet1/0/1] qos wrr 3 group sp
[Sysname-GigabitEthernet1/0/1] qos wrr 4 group 1 weight 2
[Sysname-GigabitEthernet1/0/1] qos wrr 5 group 1 weight 4
[Sysname-GigabitEthernet1/0/1] qos wrr 6 group 1 weight 6
[Sysname-GigabitEthernet1/0/1] qos wrr 7 group 1 weight 8
```

Configuring SP+WFQ queuing

Configuration procedure

To configure SP + WFQ queuing:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view or port group view.	<ul style="list-style-type: none">Enter interface view: interface <i>interface-type</i> <i>interface-number</i>Enter port group view: port-group manual <i>port-group-name</i>	Use either command. Settings in interface view take effect on the current interface. Settings in port group view take effect on all ports in the port group.
3. Enable byte-count or packet-based WFQ queuing.	qos wfq [byte-count weight]	By default, WRR queuing is enabled.
4. Configure SP queue scheduling.	qos wfq <i>queue-id</i> group sp	By default, all the queues of a WFQ-enabled port are in the WFQ group.
5. Configure the scheduling weight for a queue.	qos wfq <i>queue-id</i> group <i>group-id</i> { weight byte-count } <i>schedule-value</i>	By default, the scheduling weight is 1 for each queue of a WFQ-enabled port.
6. Configure the minimum guaranteed bandwidth for a queue.	qos bandwidth queue <i>queue-id</i> min <i>bandwidth-value</i>	Optional. 64 kbps for each queue by default.

NOTE:

To guarantee successful WFQ configuration, make sure that the scheduling weight type (byte-count or packet-based) is the same as the WFQ queuing type (byte-count or packet-based) when you configure the scheduling weight for a WFQ queue.

Configuration example

Network requirements

- Configure SP+WFQ queuing on GigabitEthernet 1/0/1, and use packet-based WFQ scheduling weights.
- Configure queue 0, queue 1, queue 2, and queue 3 on GigabitEthernet 1/0/1 to be in SP queue scheduling group.
- Configure queue 4, queue 5, queue 6, and queue 7 on GigabitEthernet 1/0/1 to use WFQ queuing, with the weight 2, 4, 6, and 8 and the minimum guaranteed bandwidth 128 kbps.

Configuration procedure

Enter system view.

```
<Sysname> system-view
```

Enable the SP+WFQ queue scheduling algorithm on GigabitEthernet1/0/1.

```
[Sysname] interface GigabitEthernet 1/0/1
[Sysname-GigabitEthernet1/0/1] qos wfq weight
[Sysname-GigabitEthernet1/0/1] qos wfq 0 group sp
[Sysname-GigabitEthernet1/0/1] qos wfq 1 group sp
[Sysname-GigabitEthernet1/0/1] qos wfq 2 group sp
[Sysname-GigabitEthernet1/0/1] qos wfq 3 group sp
[Sysname-GigabitEthernet1/0/1] qos wfq 4 group 1 weight 2
[Sysname-GigabitEthernet1/0/1] qos bandwidth queue 4 min 128
[Sysname-GigabitEthernet1/0/1] qos wfq 5 group 1 weight 4
[Sysname-GigabitEthernet1/0/1] qos bandwidth queue 5 min 128
[Sysname-GigabitEthernet1/0/1] qos wfq 6 group 1 weight 6
[Sysname-GigabitEthernet1/0/1] qos bandwidth queue 6 min 128
[Sysname-GigabitEthernet1/0/1] qos wfq 7 group 1 weight 8
[Sysname-GigabitEthernet1/0/1] qos bandwidth queue 7 min 128
```

Configuring congestion avoidance (available only on the 5500 EI)

Both bridge mode (Layer 2) and route mode (Layer 3) Ethernet ports support the congestion avoidance function. The term *interface* in this chapter collectively refers to these types of ports. You can use the **port link-mode** command to set an Ethernet port to operate in bridge or route mode (see *Layer 2—LAN Switching Configuration Guide*).

Overview

Avoiding congestion before it occurs is a proactive approach to improving network performance. As a flow control mechanism, congestion avoidance actively monitors network resources (such as queues and memory buffers), and drops packets when congestion is expected to occur or deteriorate.

Compared with end-to-end flow control, this flow control mechanism controls the load of more flows in a device. When dropping packets from a source end, it cooperates with the flow control mechanism (such as TCP flow control) at the source end to regulate the network traffic size. The combination of the local packet drop policy and the source-end flow control mechanism helps maximize throughput and network use efficiency and minimize packet loss and delay.

Tail drop

Congestion management techniques drop all packets that are arriving at a full queue. This tail drop mechanism results in global TCP synchronization. If packets from multiple TCP connections are dropped, these TCP connections go into the state of congestion avoidance and slow start to reduce traffic, but traffic peak occurs later. Consequently, the network traffic jitters all the time.

RED and WRED

You can use random early detection (RED) or weighted random early detection (WRED) to avoid global TCP synchronization.

Both RED and WRED avoid global TCP synchronization by randomly dropping packets. When the sending rates of some TCP sessions slow down after their packets are dropped, other TCP sessions remain at high sending rates. Link bandwidth is efficiently used, because TCP sessions at high sending rates always exist.

The RED or WRED algorithm sets an upper threshold and lower threshold for each queue, and processes the packets in a queue as follows:

- When the queue size is shorter than the lower threshold, no packet is dropped;
- When the queue size reaches the upper threshold, all subsequent packets are dropped;
- When the queue size is between the lower threshold and the upper threshold, the received packets are dropped at random. The drop probability in a queue increases along with the queue size under the maximum drop probability.

NOTE:

The Switch Series does not support the upper threshold configuration.

Introduction to WRED configuration

On the 5500 EI Switch Series, WRED is implemented with WRED tables. WRED tables are created globally in system view and then applied to interfaces.

Before configuring WRED, determine the following parameters:

- **Lower threshold**—When the average queue length is below the lower threshold, no packet is dropped. When the average queue length exceeds the lower threshold, the switch drops packets at the user-configured drop probability.
- **Drop precedence**—A parameter used in packet drop. Value 0 represents green packets, 1 represents yellow packets, and 2 represents red packets. Red packets are preferentially dropped.
- **Denominator**—Denominator for drop probability calculation. A greater denominator means a lower drop probability. Table 4 shows the denominator and the drop probability dependencies.

Table 4 Denominator and the drop probability dependencies

Denominator	Drop probability
0	100%
1 to 8	1/8
9 to 16	1/16
17 to 32	1/32
33 to 64	1/64
65 to 128	1/128

In a WRED table, drop parameters are configured on a per queue basis because WRED regulates packets on a per queue basis.

A WRED table can be applied to multiple interfaces. For a WRED table already applied to an interface, you can modify the values of the WRED table, but you cannot remove the WRED table.

Configuration procedure

To configure and apply a queue-based WRED table:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a WRED table and enter its view.	qos wred queue table <i>table-name</i>	N/A
3. Configure the other WRED parameters.	queue <i>queue-value</i> [drop-level <i>drop-level</i>] low-limit <i>low-limit</i> [discard-probability <i>discard-prob</i>]	Optional. By default, <i>low-limit</i> is 10, and <i>discard-prob</i> is 10.

Step	Command	Remarks
4. Enter interface view or port group view.	<ul style="list-style-type: none"> Enter interface view: interface <i>interface-type</i> <i>interface-number</i> Enter port group view: port-group manual <i>port-group-name</i> 	Use either command. Settings in interface view take effect on the current interface. Settings in port group view take effect on all ports in the port group.
5. Apply the WRED table to the interface or port group.	qos wred apply <i>table-name</i>	N/A

Configuration examples

Apply a WRED table to Layer 2 port GigabitEthernet 1/0/1. Set the *low-limit* to 30 and *discard-prob* to 20 for queue 1.

Enter system view.

```
<Sysname> system-view
```

Create a queue-based WRED table named **queue-table1**, and configure the drop parameters.

```
[Sysname] qos wred queue table queue-table1
```

```
[Sysname-wred-table-queue-table1] queue 1 low-limit 30 discard-probability 20
```

```
[Sysname-wred-table-queue-table1] quit
```

Enter port view.

```
[Sysname] interface gigabitethernet 1/0/1
```

Apply the WRED table to GigabitEthernet 1/0/1.

```
[Sysname-GigabitEthernet1/0/1] qos wred apply queue-table1
```

Displaying and maintaining WRED

Task	Command	Remarks
Display WRED configuration information on the interface or all interfaces.	display qos wred interface [<i>interface-type</i> <i>interface-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display configuration information about a WRED table or all WRED tables.	display qos wred table [<i>table-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Configuring traffic filtering

Traffic filtering filters traffic matching certain criteria. For example, you can filter packets sourced from a specific IP address according to network status.

Configuration procedure

To configure traffic filtering:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a class and enter class view.	traffic classifier <i>tcl-name</i> [operator { and or }]	N/A
3. Configure match criteria.	if-match <i>match-criteria</i>	N/A
4. Return to system view.	quit	N/A
5. Create a behavior and enter behavior view.	traffic behavior <i>behavior-name</i>	N/A
6. Configure the traffic filtering action.	filter { deny permit }	<ul style="list-style-type: none">• deny—Drops packets.• permit—Permits packets to pass through.
7. Return to system view.	quit	N/A
8. Create a policy and enter policy view.	qos policy <i>policy-name</i>	N/A
9. Associate the class with the traffic behavior in the QoS policy.	classifier <i>tcl-name</i> behavior <i>behavior-name</i>	N/A
10. Return to system view.	quit	N/A
11. Apply the QoS policy.	<ul style="list-style-type: none">• Applying the QoS policy to an interface• Applying the QoS policy to online users• Applying the QoS policy to a VLAN• Applying the QoS policy globally• Applying the QoS policy to the control plane	Choose one application destination as needed.
12. Display the traffic filtering configuration.	display traffic behavior user-defined [<i>behavior-name</i>] [[{ begin exclude include } <i>regular-expression</i>]	Optional. Available in any view

NOTE:

With **filter deny** configured for a traffic behavior, the other actions (except class-based accounting and traffic mirroring) in the traffic behavior do not take effect.

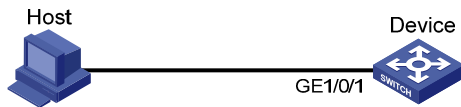
Traffic filtering configuration example

Network requirements

As shown in [Figure 18](#), Host is connected to GigabitEthernet 1/0/1 of Device.

Configure traffic filtering to filter the packets with source port being 21, and received on GigabitEthernet 1/0/1.

Figure 18 Network diagram



Configuration procedure

Create advanced ACL 3000, and configure a rule to match packets whose source port number is 21.

```
<DeviceA> system-view
[DeviceA] acl number 3000
[DeviceA-acl-adv-3000] rule 0 permit tcp source-port eq 21
[DeviceA-acl-adv-3000] quit
```

Create a class named **classifier_1**, and use ACL 3000 as the match criterion in the class.

```
[DeviceA] traffic classifier classifier_1
[DeviceA-classifier-classifier_1] if-match acl 3000
[DeviceA-classifier-classifier_1] quit
```

Create a behavior named **behavior_1**, and configure the traffic filtering action to drop packets.

```
[DeviceA] traffic behavior behavior_1
[DeviceA-behavior-behavior_1] filter deny
[DeviceA-behavior-behavior_1] quit
```

Create a policy named **policy**, and associate class **classifier_1** with behavior **behavior_1** in the policy.

```
[DeviceA] qos policy policy
[DeviceA-qospolicy-policy] classifier classifier_1 behavior behavior_1
[DeviceA-qospolicy-policy] quit
```

Apply the policy named **policy** to the incoming traffic of GigabitEthernet 1/0/1.

```
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] qos apply policy policy inbound
```

Configuring priority marking

Priority marking sets the priority fields or flag bits of packets to modify the priority of traffic. For example, you can use priority marking to set IP precedence or DSCP for a class of IP traffic to change its transmission priority in the network.

Priority marking can be used together with priority mapping. For more information about priority mapping, see "[Configuring priority mapping](#)."

Color-based priority marking

Coloring a packet

The switch colors a packet to indicate its transmission priority after evaluating the status of processing resources and the priority of the packet.

The switch can color a packet by using one of the following approaches:

- Uses the token bucket mechanism (bucket C and bucket E) of traffic policing:
 - If bucket C has enough tokens, the packet is colored green.
 - If bucket C does not have enough tokens but bucket E has enough tokens, the packet is colored yellow.
 - If neither bucket C nor bucket E has enough tokens, the packet is colored red.
- If traffic policing is not configured, looks up the 802.1p priority of a packet in the 802.1p-to-drop priority mapping table, allocates drop precedence to the packet, and colors the packet according to the drop precedence.
 - Drop precedence 0 represents green packets.
 - Drop precedence 1 represents yellow packets.
 - Drop precedence 2 represents red packets.

For more information about traffic policing, see "[Configuring traffic policing, traffic shaping, and line rate](#)." For more information about priority mapping tables, see "[Configuring priority mapping](#)."

Marking packets based on their colors

Color-based priority marking supports re-marking DSCP precedence.

You can configure color-based marking in the following ways:

- To mark packets based on a color set during traffic policing, configure a priority marking action for the color in the traffic policing action **car**. For more information, see "[Configuring traffic policing](#)."
- To mark packets based on their drop precedence, configure a priority marking action for a color by using the **remark** command as described in the subsequent section.

ⓘ IMPORTANT:

Do not use the **remark** command together with the **car** command in a traffic behavior to perform color-based marking.

Configuration procedure

To configure priority marking:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a class and enter class view.	traffic classifier <i>tcl-name</i> [operator { and or }]	N/A
3. Configure match criteria.	if-match <i>match-criteria</i>	N/A
4. Return to system view.	quit	N/A
5. Create a behavior and enter behavior view.	traffic behavior <i>behavior-name</i>	N/A
6. Set the DSCP value for packets.	remark [green red yellow] dscp <i>dscp-value</i>	Optional.
7. Set the 802.1p priority for packets or configure the inner-to-outer tag priority copying function.	remark dot1p { <i>8021p</i> customer-dot1p-trust }	Optional.
8. Set the drop precedence for packets.	remark drop-precedence <i>drop-precedence-value</i>	Optional. Applicable to only the outbound direction.
9. Set the IP precedence for packets.	remark ip-precedence <i>ip-precedence-value</i>	Optional.
10. Set the local precedence for packets.	remark local-precedence <i>local-precedence</i>	Optional.
11. Return to system view.	quit	N/A
12. Create a policy and enter policy view.	qos policy <i>policy-name</i>	N/A
13. Associate the class with the traffic behavior in the QoS policy.	classifier <i>tcl-name</i> behavior <i>behavior-name</i>	N/A
14. Return to system view.	quit	N/A
15. Apply the QoS policy.	<ul style="list-style-type: none"> • Applying the QoS policy to an interface • Applying the QoS policy to online users • Applying the QoS policy to a VLAN • Applying the QoS policy globally • Applying the QoS policy to the control plane 	Choose one application destination as needed.
16. Display the priority marking configuration.	display traffic behavior user-defined [<i>behavior-name</i>] [[{ begin exclude include } <i>regular-expression</i>]	Optional. Available in any view

The following table shows the support for priority marking actions in the inbound and outbound directions.

Table 5 Support for priority marking actions in the inbound and outbound directions

Action	inbound	outbound
802.1p priority marking	Yes	Yes
Drop precedence marking	Yes	No
DSCP marking	Yes	Yes
IP precedence marking	Yes	Yes
Local precedence marking	Yes	No

Local precedence re-marking configuration example

Network requirements

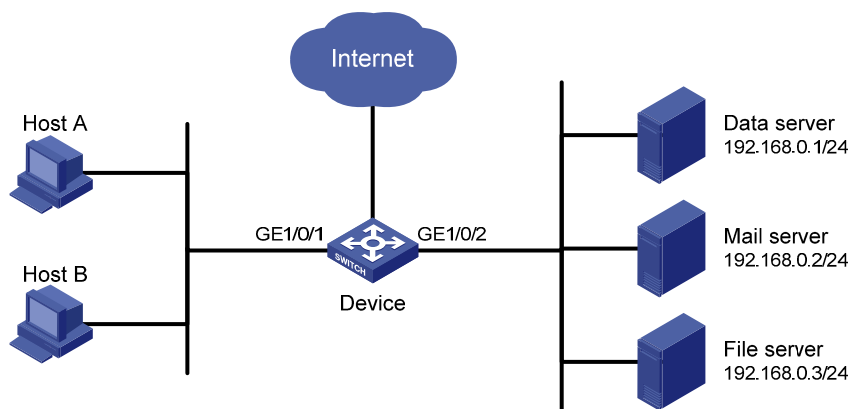
As shown in [Figure 19](#), the company's enterprise network interconnects hosts with servers through Device. The network is described as follows:

- Host A and Host B are connected to GigabitEthernet 1/0/1 of Device.
- The data server, mail server, and file server are connected to GigabitEthernet 1/0/2 of Device.

Configure priority marking on Device to satisfy the following requirements:

Traffic source	Destination	Processing priority
Host A, B	Data server	High
Host A, B	Mail server	Medium
Host A, B	File server	Low

Figure 19 Network diagram



Configuration procedure

Create advanced ACL 3000, and configure a rule to match packets with destination IP address 192.168.0.1.

```

<Device> system-view
[Device] acl number 3000
[Device-acl-adv-3000] rule permit ip destination 192.168.0.1 0
[Device-acl-adv-3000] quit

# Create advanced ACL 3001, and configure a rule to match packets with destination IP address
192.168.0.2.
[Device] acl number 3001
[Device-acl-adv-3001] rule permit ip destination 192.168.0.2 0
[Device-acl-adv-3001] quit

# Create advanced ACL 3002, and configure a rule to match packets with destination IP address
192.168.0.3.
[Device] acl number 3002
[Device-acl-adv-3002] rule permit ip destination 192.168.0.3 0
[Device-acl-adv-3002] quit

# Create a class named classifier_dbserver, and use ACL 3000 as the match criterion in the class.
[Device] traffic classifier classifier_dbserver
[Device-classifier-classifier_dbserver] if-match acl 3000
[Device-classifier-classifier_dbserver] quit

# Create a class named classifier_mserver, and use ACL 3001 as the match criterion in the class.
[Device] traffic classifier classifier_mserver
[Device-classifier-classifier_mserver] if-match acl 3001
[Device-classifier-classifier_mserver] quit

# Create a class named classifier_fserver, and use ACL 3002 as the match criterion in the class.
[Device] traffic classifier classifier_fserver
[Device-classifier-classifier_fserver] if-match acl 3002
[Device-classifier-classifier_fserver] quit

# Create a behavior named behavior_dbserver, and configure the action of setting the local precedence
value to 4.
[Device] traffic behavior behavior_dbserver
[Device-behavior-behavior_dbserver] remark local-precedence 4
[Device-behavior-behavior_dbserver] quit

# Create a behavior named behavior_mserver, and configure the action of setting the local precedence
value to 3.
[Device] traffic behavior behavior_mserver
[Device-behavior-behavior_mserver] remark local-precedence 3
[Device-behavior-behavior_mserver] quit

# Create a behavior named behavior_fserver, and configure the action of setting the local precedence
value to 2.
[Device] traffic behavior behavior_fserver
[Device-behavior-behavior_fserver] remark local-precedence 2
[Device-behavior-behavior_fserver] quit

# Create a policy named policy_server, and associate classes with behaviors in the policy.
[Device] qos policy policy_server
[Device-qospolicy-policy_server] classifier classifier_dbserver behavior
behavior_dbserver

```

```
[Device-qospolicy-policy_server] classifier classifier_mserver behavior
behavior_mserver
[Device-qospolicy-policy_server] classifier classifier_fserver behavior
behavior_fserver
[Device-qospolicy-policy_server] quit
# Apply the policy named policy_server to the incoming traffic of GigabitEthernet 1/0/1.
[Device] interface gigabitethernet 1/0/1
[Device-GigabitEthernet1/0/1] qos apply policy policy_server inbound
[Device-GigabitEthernet1/0/1] quit
```

Configuring traffic redirecting

Traffic redirecting is the action of redirecting the packets matching the specific match criteria to a certain location for processing.

The following redirect actions are supported:

- **Redirecting traffic to the CPU**—redirects packets that require processing by the CPU to the CPU.
- **Redirecting traffic to an interface**—redirects packets that require processing by an interface to the interface. Note that this action applies to only Layer 2 packets, and the target interface must be a Layer 2 interface.
- **Redirecting traffic to the next hop**—redirects packets that require processing by an interface to the interface. This action only applies to Layer 3 packets.

❗ **IMPORTANT:**

The 5500 SI switch does not support redirecting traffic to the next hop.

Configuration restrictions and guidelines

- The actions of redirecting traffic to the CPU, redirecting traffic to an interface, and redirecting traffic to the next hop are mutually exclusive with each other in the same traffic behavior.
- A QoS policy with traffic redirecting actions can be applied to only the inbound direction of a port, VLAN, or all ports.
- The default of the **fail-action** keyword is **forward**.
- You can use the **display traffic behavior user-defined** command to view the traffic redirecting configuration.

Configuration procedure

To configure traffic redirecting:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a class and enter class view.	traffic classifier <i>tcl-name</i> [operator { and or }]	N/A
3. Configure match criteria.	if-match <i>match-criteria</i>	N/A
4. Return to system view.	quit	N/A
5. Create a behavior and enter behavior view.	traffic behavior <i>behavior-name</i>	N/A

Step	Command	Remarks
6. Configure a traffic redirecting action.	redirect { cpu interface <i>interface-type interface-number</i> next-hop { <i>ipv4-add1</i> [<i>ipv4-add2</i>] <i>ipv6-add1</i> [<i>interface-type interface-number</i>] [<i>ipv6-add2</i> [<i>interface-type interface-number</i>]] } [fail-action { discard forward }] }	The 5500 SI switch does not support the next-hop or fail-action keyword.
7. Return to system view.	quit	N/A
8. Create a policy and enter policy view.	qos policy <i>policy-name</i>	N/A
9. Associate the class with the traffic behavior in the QoS policy.	classifier <i>tcl-name</i> behavior <i>behavior-name</i>	N/A
10. Return to system view.	quit	N/A
11. Apply the QoS policy.	<ul style="list-style-type: none"> • Applying the QoS policy to an interface • Applying the QoS policy to a VLAN • Applying the QoS policy globally • Applying the QoS policy to the control plane 	Choose one application destination as needed.

Redirect-to-next hop configuration example

Network requirements

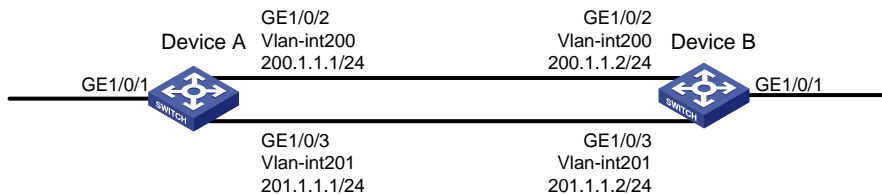
As shown in [Figure 20](#), the network is described as follows:

- Device A is connected to Device through two links. At the same time, Device A and Device B are each connected to other devices.
- GigabitEthernet 1/0/2 of Device A and GigabitEthernet 1/0/2 of Device B belong to VLAN 200.
- Ethernet 1/3 of Device A and Ethernet 1/3 of Device B belong to VLAN 201.
- On Device A, the IP address of VLAN-interface 200 is 200.1.1.1/24, and that of VLAN-interface 201 is 201.1.1.1/24.
- On Device B, the IP address of VLAN-interface 200 is 200.1.1.2/24, and that of VLAN-interface 201 is 201.1.1.2/24.

Configure the actions of redirecting traffic to the next hop to implement policy-based routing and satisfy the following requirements:

- Packets with source IP address 2.1.1.1 received on GigabitEthernet 1/0/1 of Device A are forwarded to IP address 200.1.1.2.
- Packets with source IP address 2.1.1.2 received on GigabitEthernet 1/0/1 of Device A are forwarded to IP address 201.1.1.2.
- Other packets received on Ethernet 1/1 of Device A are forwarded according to the routing table.

Figure 20 Network diagram



Configuration procedure

Create basic ACL 2000, and configure a rule to match packets with source IP address 2.1.1.1.

```
<DeviceA> system-view
[DeviceA] acl number 2000
[DeviceA-acl-basic-2000] rule permit source 2.1.1.1 0
[DeviceA-acl-basic-2000] quit
```

Create basic ACL 2001, and configure a rule to match packets with source IP address 2.1.1.2.

```
[DeviceA] acl number 2001
[DeviceA-acl-basic-2001] rule permit source 2.1.1.2 0
[DeviceA-acl-basic-2001] quit
```

Create a class named **classifier_1**, and use ACL 2000 as the match criterion in the class.

```
[DeviceA] traffic classifier classifier_1
[DeviceA-classifier-classifier_1] if-match acl 2000
[DeviceA-classifier-classifier_1] quit
```

Create a class named **classifier_2**, and use ACL 2001 as the match criterion in the class.

```
[DeviceA] traffic classifier classifier_2
[DeviceA-classifier-classifier_2] if-match acl 2001
[DeviceA-classifier-classifier_2] quit
```

Create a behavior named **behavior_1**, and configure the action of redirecting traffic to the next hop 200.1.1.2.

```
[DeviceA] traffic behavior behavior_1
[DeviceA-behavior-behavior_1] redirect next-hop 200.1.1.2
[DeviceA-behavior-behavior_1] quit
```

Create a behavior named **behavior_2**, and configure the action of redirecting traffic to the next hop 200.1.1.2.

```
[DeviceA] traffic behavior behavior_2
[DeviceA-behavior-behavior_2] redirect next-hop 201.1.1.2
[DeviceA-behavior-behavior_2] quit
```

Create a policy named **policy**, associate class **classifier_1** with behavior **behavior_1**, and associate class **classifier_2** with behavior **behavior_2** in the policy.

```
[DeviceA] qos policy policy
[DeviceA-qospolicy-policy] classifier classifier_1 behavior behavior_1
[DeviceA-qospolicy-policy] classifier classifier_2 behavior behavior_2
[DeviceA-qospolicy-policy] quit
```

Apply the policy named **policy** to the incoming traffic of GigabitEthernet 1/0/1.

```
[DeviceA] interface gigabitethernet 1/0/1
```



```
[DeviceA-GigabitEthernet1/0/1] qos apply policy policy inbound
```

Configuring class-based accounting

Class-based accounting collects statistics (in packets or bytes) on a per-traffic class basis. For example, you can define the action to collect statistics for traffic sourced from a certain IP address. By analyzing the statistics, you can determine whether anomalies have occurred and what action to take. The 5500 EI and 5500 SI Switch Series supports only collecting statistics in packets.

Configuration procedure

To configure class-based accounting:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a class and enter class view.	traffic classifier <i>tcl-name</i> [operator { and or }]	N/A
3. Configure match criteria.	if-match <i>match-criteria</i>	N/A
4. Return to system view.	quit	N/A
5. Create a behavior and enter behavior view.	traffic behavior <i>behavior-name</i>	N/A
6. Configure the accounting action.	accounting	N/A
7. Return to system view.	quit	N/A
8. Create a policy and enter policy view.	qos policy <i>policy-name</i>	N/A
9. Associate the class with the traffic behavior in the QoS policy.	classifier <i>tcl-name</i> behavior <i>behavior-name</i>	N/A
10. Return to system view.	quit	N/A
11. Apply the QoS policy.	<ul style="list-style-type: none">• Applying the QoS policy to an interface• Applying the QoS policy to a VLAN• Applying the QoS policy globally• Applying the QoS policy to the control plane	Choose one application destination as needed.

Displaying and maintaining traffic accounting

You can verify the configuration with the **display qos policy global**, **display qos policy interface**, or **display qos vlan-policy** command depending on the occasion where the QoS policy is applied.

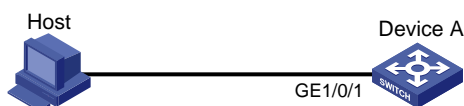
Class-based accounting configuration example

Network requirements

As shown in [Figure 21](#), Host is connected to GigabitEthernet 1/0/1 of Device A.

Configure class-based accounting to collect statistics for traffic sourced from 1.1.1.1/24 and received on GigabitEthernet 1/0/1.

Figure 21 Network diagram



Configuration procedure

Create basic ACL 2000, and configure a rule to match packets with source IP address 1.1.1.1.

```
<DeviceA> system-view
[DeviceA] acl number 2000
[DeviceA-acl-basic-2000] rule permit source 1.1.1.1 0
[DeviceA-acl-basic-2000] quit
```

Create a class named **classifier_1**, and use ACL 2000 as the match criterion in the class.

```
[DeviceA] traffic classifier classifier_1
[DeviceA-classifier-classifier_1] if-match acl 2000
[DeviceA-classifier-classifier_1] quit
```

Create a behavior named **behavior_1**, and configure the traffic accounting action.

```
[DeviceA] traffic behavior behavior_1
[DeviceA-behavior-behavior_1] accounting
[DeviceA-behavior-behavior_1] quit
```

Create a policy named **policy**, and associate class **classifier_1** with behavior **behavior_1** in the policy.

```
[DeviceA] qos policy policy
[DeviceA-qospolicy-policy] classifier classifier_1 behavior behavior_1
[DeviceA-qospolicy-policy] quit
```

Apply the policy named **policy** to the incoming traffic of GigabitEthernet 1/0/1.

```
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] qos apply policy policy inbound
[DeviceA-GigabitEthernet1/0/1] quit
```

Display traffic statistics to verify the configuration.

```
[DeviceA] display qos policy interface gigabitethernet 1/0/1
```

```
Interface: GigabitEthernet1/0/1
```

```
Direction: Inbound
```

```
Policy: policy
```

```
Classifier: classifier_1
```

Operator: AND
Rule(s) : If-match acl 2000
Behavior: behavior_1
Accounting Enable:
28529 (Packets)

Configuring the data buffer

Overview

Data buffer

The Switch Series provides the data buffer to buffer packets to be sent out ports to avoid packet loss when bursty traffic causes congestion.

The switch controls how a port uses the data buffer by allocating the cell resource and packet resource (called "buffer resources").

- The cell resource is the physical storage space in cells for the data buffer. The cell resource allocated to a port indicates the maximum buffer space that the port can occupy in the buffer.
- The packet resource is the logical buffering space in packets. A packet is counted as one regardless of its length. A packet in the packet resource uses a certain amount of cell resources in the cell resource, depending on its length. The packet resource allocated to a port indicates the maximum number of packets that the port can store in the buffer.

Set independently, the packet resource and the cell resource work simultaneously to regulate data buffering. A packet can be buffered only when both resources are adequate.

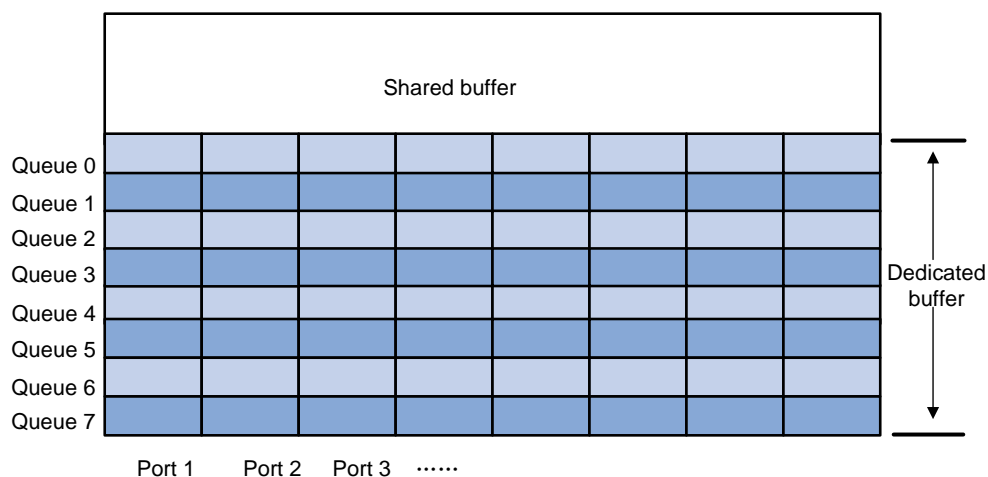
Data buffer allocation

Cell resource allocation

The cell resource is divided into a shared resource and a dedicated resource. You can manually set the percentage of the shared resource to the total buffer, and the remaining buffer becomes the dedicated resource automatically.

On a 5500 EI and 5500 SI switch, the cell resource is allocated as shown in [Figure 22](#).

Figure 22 Cell resource allocation on the 5500 EI and 5500 SI



The dedicated resource is allocated following these rules:

- **On a per-port basis**—As illustrated by the vertical lines in [Figure 22](#), the switch automatically divides the dedicated resource among all ports evenly.
- **On a per-queue basis**—As illustrated by the horizontal lines in [Figure 22](#), the dedicated resource of each port is proportionately allocated among the queues on it and all ports use the same allocation scheme. The percentage of the resource allocated to a queue is called the minimum guaranteed resource percentage of the queue.

The shared buffer in the cell resource can buffer the bursty traffic on ports. The shared resource is shared by all queues of all ports. When a certain queue of a port is congested because its dedicated cell resource gets full, it can use a certain portion of the shared resource of the cell resource. The maximum shared resource size available for a queue is defined as a percentage of the shared resource. After the bursty traffic is transmitted, the shared resource used by the bursty traffic is released for other ports or queues to use. For example, you can configure port 1 to use 30% of the shared buffer of the cell resource.

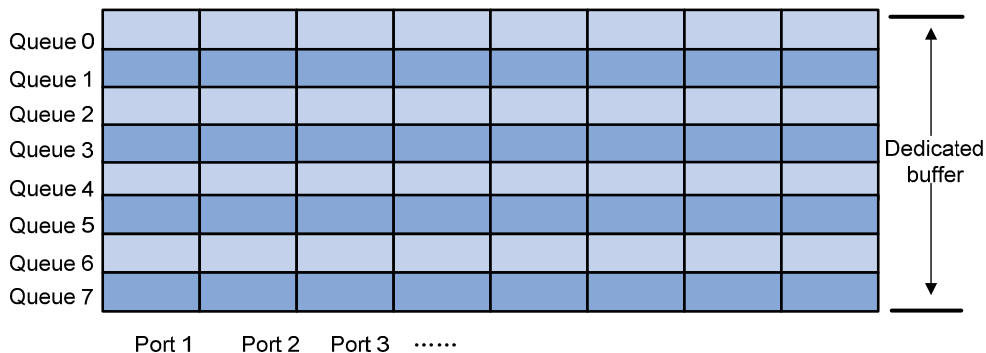
You can perform the following parameters for the cell resource:

- Configure the shared resource size
- Configure the minimum guaranteed resource size for a queue
- Configure the maximum shared resource size for a port

Packet resource allocation

Different from the cell resource, the packet resource does not have a shared resource, and the whole buffer is evenly allocated to ports, as shown in [Figure 23](#).

Figure 23 Packet resource allocation on the 5500 EI and 5500 SI



Like the packet resource, the cell resource is also allocated on a per-port basis and on a per-queue basis. Because the packet resource does not have the shared resource, you can adjust only the percentage of the queue buffer to the port buffer, which is called the minimum guaranteed buffer percentage.

Data buffer configuration approaches

You can configure the data buffer on the 5500 EI and 5500 SI Switch Series in one of the following approaches:

- [Using the burst function to configure the data buffer setup](#)
- [Manually configuring the data buffer setup](#)

NOTE:

The two approaches are mutually exclusive. If the data buffer setup has been configured in one approach, you must remove the present configuration first before you use the other approach.

Using the burst function to configure the data buffer setup

The burst function allows the switch to automatically determine the shared resource size, the minimum guaranteed resource size for each queue, the maximum shared resource size for each queue, and the maximum shared resource size per port.

The burst function helps optimize packet buffering to ameliorate forwarding performance in the following scenarios:

- Broadcast or multicast traffic is dense and bursts of traffic are usually large.
- High-speed traffic is forwarded over low-speed links or traffic received from multiple ports is forwarded through a port operating at the same speed.

To use the burst function to configure the data buffer:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable the burst function.	burst-mode enable	By default, the burst function is disabled.

Manually configuring the data buffer setup

Data buffer configuration is complicated and has significant impacts on the forwarding performance of a device. HP does not recommend modifying the data buffer parameters unless you are sure that your device will benefit from the change. If a larger buffer is needed, HP recommends that you enable the burst function to allocate the buffer automatically.

Manually configuring the data buffer

Complete the following tasks to manually configure the data buffer:

Task	Remarks	
Configuring the cell resource	Configuring the shared resource size	Optional
	Configuring the minimum guaranteed resource size for a queue	Optional
	Configuring the maximum shared resource size for a port	Optional
Configuring the packet resource	Configuring the minimum guaranteed resource size for a queue	Optional
Applying the data buffer settings	Required	

Configuring the cell resource

Configuring the shared resource size

To configure the shared resource size:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the shared resource area of the cell resource in percentage.	buffer egress [slot <i>slot-number</i>] cell total-shared ratio <i>ratio</i>	Optional. By default, the shared resource area of the cell resource is 60%.

Configuring the minimum guaranteed resource size for a queue

When configuring the minimum guaranteed resource size for a queue, follow these guidelines:

- Modifying the minimum guaranteed resource size for a queue can affect those of the other queues, because the dedicated resource of a port is shared by eight queues. The system will automatically allocate the remaining dedicated resource space among all queues that are not manually assigned a minimum guaranteed resource space. For example, if you set the minimum guaranteed resource size to 30% for a queue, the remaining seven queues will each share 10% of the dedicated resource of the port.
- The minimum guaranteed resource settings of a queue apply to the queue with the same number on each port.

To configure the minimum guaranteed resource size for a queue:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the minimum guaranteed cell resource size for a queue as a percentage of the dedicated cell resource per port.	buffer egress [slot <i>slot-number</i>] cell queue <i>queue-id</i> guaranteed ratio <i>ratio</i>	Optional. By default, the minimum guaranteed resource size for a queue is 12% of the dedicated resource of the port in the cell resource.

Configuring the maximum shared resource size for a port

To configure the maximum shared resource size for a port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the maximum shared cell resource size for a queue as a percentage of the shared cell resource.	buffer egress [slot <i>slot-number</i>] cell shared ratio <i>ratio</i>	Optional. By default, a queue can use up to 50% of the shared cell resource.

NOTE:

The maximum shared resource settings for a queue apply to the queue with the same number on each port.

Configuring the packet resource

Configuring the minimum guaranteed resource size for a queue

When configuring the minimum guaranteed resource size for a queue, follow these guidelines:

- Modifying the minimum guaranteed resource size for a queue can affect those of the other queues, because the dedicated resource of a port is shared by eight queues. The system will automatically allocate the remaining dedicated resource space among all queues that are not manually assigned a minimum guaranteed resource space. For example, if you set the minimum guaranteed resource size to 30% for a queue, the remaining seven queues will each share 10% of the dedicated resource of the port.
- The minimum guaranteed resource settings of a queue apply to the queue with the same number on each port.

To configure the minimum guaranteed resource size for a queue:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the minimum guaranteed packet resource size for a queue as a percentage of the dedicated packet resource per port.	buffer egress [slot <i>slot-number</i>] packet queue <i>queue-id</i> guaranteed ratio <i>ratio</i>	Optional. By default, the minimum guaranteed resource size for queue 2 is 51% of the dedicated resource of the port in the packet resource, and that for any other queue is 7%.

Applying the data buffer settings

After manually configuring data buffer, you should execute the following steps to make the manual data buffer configurations take effect.

To apply the data buffer settings:

Step	Command
1. Enter system view.	system-view
2. Apply the data buffer settings.	buffer apply

Appendix A Default priority mapping tables

Uncolored priority mapping tables

For the default **dscp-dscp** mapping table, an input value yields a target value equal to it.

Table 6 Default dot1p-lp and dot1p-dp priority mapping tables

Input priority value	dot1p-lp mapping	dot1p-dp mapping
802.1p priority (dot1p)	Local precedence (lp)	Drop precedence (dp)
0	2	0
1	0	0
2	1	0
3	3	0
4	4	0
5	5	0
6	6	0
7	7	0

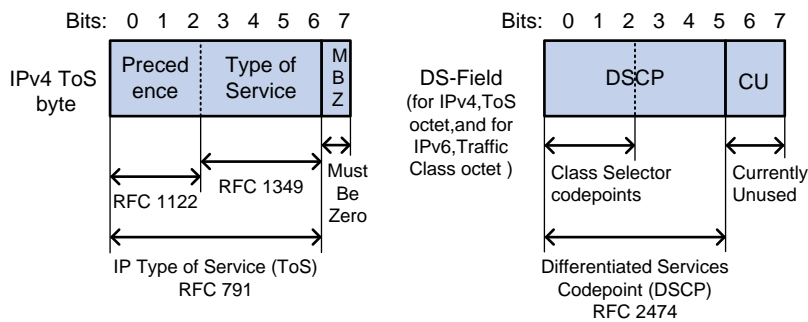
Table 7 Default dscp-dp and dscp-dot1p priority mapping tables

Input priority value	dscp-dp mapping	dscp-dot1p mapping
DSCP	Drop precedence (dp)	802.1p priority (dot1p)
0 to 7	0	0
8 to 15	0	1
16 to 23	0	2
24 to 31	0	3
32 to 39	0	4
40 to 47	0	5
48 to 55	0	6
56 to 63	0	7

Appendix B Packet precedences

IP precedence and DSCP values

Figure 24 ToS and DS fields



As shown in Figure 24, the ToS field in the IPv4 header contains eight bits, where the first three bits (0 to 2) represent IP precedence from 0 to 7; the Traffic Classes field in the IPv6 header contains eight bits, where the first three bits (0 to 2) represent IP precedence from 0 to 7. According to RFC 2474, the ToS field in the IPv4 header or the Traffic Classes field in the IPv6 header is redefined as the differentiated services (DS) field, where a DSCP value is represented by the first six bits (0 to 5) and is in the range 0 to 63. The remaining two bits (6 and 7) are reserved.

Table 8 Description on IP precedence

IP precedence (decimal)	IP precedence (binary)	Description
0	000	Routine
1	001	priority
2	010	immediate
3	011	flash
4	100	flash-override
5	101	critical
6	110	internet
7	111	network

Table 9 Description on DSCP values

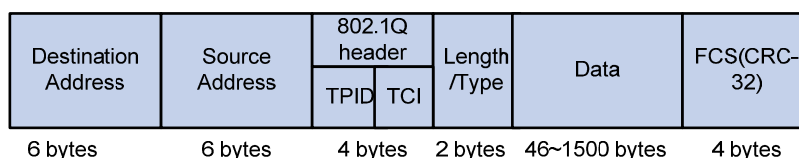
DSCP value (decimal)	DSCP value (binary)	Description
46	101110	ef
10	001010	af11
12	001100	af12
14	001110	af13
18	010010	af21

DSCP value (decimal)	DSCP value (binary)	Description
20	010100	af22
22	010110	af23
26	011010	af31
28	011100	af32
30	011110	af33
34	100010	af41
36	100100	af42
38	100110	af43
8	001000	cs1
16	010000	cs2
24	011000	cs3
32	100000	cs4
40	101000	cs5
48	110000	cs6
56	111000	cs7
0	000000	be (default)

802.1p priority

802.1p priority lies in the Layer 2 header and applies to occasions where Layer 3 header analysis is not needed and QoS must be assured at Layer 2.

Figure 25 An Ethernet frame with an 802.1Q tag header



As shown in [Figure 25](#), the four-byte 802.1Q tag header consists of the tag protocol identifier (TPID, two bytes in length), whose value is 0x8100, and the tag control information (TCI, two bytes in length). [Figure 26](#) shows the format of the 802.1Q tag header. The Priority field in the 802.1Q tag header is called the "802.1p priority", because its use is defined in IEEE 802.1p. [Table 10](#) shows the values for 802.1p priority.

Figure 26 802.1Q tag header

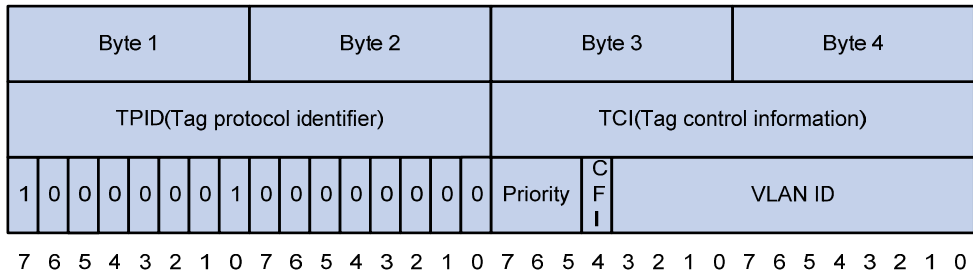


Table 10 Description on 802.1p priority

802.1p priority (decimal)	802.1p priority (binary)	Description
0	000	best-effort
1	001	background
2	010	spare
3	011	excellent-effort
4	100	controlled-load
5	101	video
6	110	voice
7	111	network-management

Index

[A](#) [C](#) [D](#) [I](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [T](#) [U](#)

A

ACL configuration task list, [4](#)

Applying the QoS policy, [22](#)

C

Changing the port priority of an interface, [30](#)

Class-based accounting configuration example, [70](#)

Color-based priority marking, [60](#)

Configuration example of using ACL for device management, [12](#)

Configuration guidelines, [29](#)

Configuration procedure, [69](#)

Configuration procedure, [65](#)

Configuration procedure, [58](#)

Configuration procedure, [61](#)

Configuration restrictions and guidelines, [65](#)

Configuring a basic ACL, [5](#)

Configuring a port to trust packet priority for priority mapping, [30](#)

Configuring a priority mapping table, [29](#)

Configuring a time range, [4](#)

Configuring an advanced ACL, [6](#)

Configuring an Ethernet frame header ACL, [9](#)

Configuring GTS, [40](#)

Configuring SP queuing, [48](#)

Configuring SP+WFQ queuing, [53](#)

Configuring SP+WRR queuing, [52](#)

Configuring the line rate, [40](#)

Configuring traffic policing, [39](#)

Configuring WFQ queuing, [50](#)

Configuring WRR queuing, [49](#)

Congestion management techniques, [44](#)

Copying an ACL, [9](#)

D

Data buffer configuration approaches, [73](#)

Defining a class, [19](#)

Defining a policy, [22](#)

Defining a traffic behavior, [21](#)

Displaying and maintaining ACLs, [11](#)

Displaying and maintaining QoS policies, [25](#)

Displaying and maintaining traffic accounting, [69](#)

Displaying and maintaining traffic policing, GTS, and line rate, [41](#)

Displaying and maintaining WRED, [57](#)

Displaying priority mappings, [31](#)

I

Introduction to WRED configuration, [56](#)

IP precedence and DSCP values, [78](#)

IPv4 packet filtering configuration example, [13](#)

IPv6 packet filtering configuration example, [14](#)

L

Local precedence re-marking configuration example, [62](#)

M

Manually configuring the data buffer setup, [74](#)

MQC approach, [18](#)

N

Non-MQC approach, [18](#)

O

Overview, [27](#)

Overview, [72](#)

Overview, [35](#)

Overview, [1](#)

Overview, [44](#)

Overview, [55](#)

Overview, [19](#)

P

Packet filtering with ACLs, [10](#)

Priority mapping table and priority marking configuration example, [32](#)

Priority trust mode configuration example, [31](#)

Q

QoS service models, [16](#)

QoS techniques, [17](#)

R

Redirect-to-next hop configuration example, [66](#)

T

Traffic filtering configuration example, [59](#)

Traffic policing configuration example, [41](#)

U

Uncolored priority mapping tables, [77](#)

Using the burst function to configure the data buffer setup, [74](#)

Contents

Configuring AAA	1
AAA overview	1
RADIUS	2
HWTACACS	7
Domain-based user management	9
RADIUS server feature of the switch	10
AAA for MPLS L3VPNs (available only on the HP 5500 EI)	11
Protocols and standards	11
RADIUS attributes	12
AAA configuration considerations and task list	15
Configuring AAA schemes	16
Configuring local users	16
Configuring RADIUS schemes	21
Configuring HWTACACS schemes	34
Configuring AAA methods for ISP domains	40
Configuration prerequisites	41
Creating an ISP domain	41
Configuring ISP domain attributes	41
Configuring AAA authentication methods for an ISP domain	42
Configuring AAA authorization methods for an ISP domain	44
Configuring AAA accounting methods for an ISP domain	46
Tearing down user connections	47
Configuring a NAS ID-VLAN binding	47
Specifying the device ID used in stateful failover mode (available only on the HP 5500 EI)	48
Configuring a switch as a RADIUS server	48
RADIUS server functions configuration task list	48
Configuring a RADIUS user	48
Specifying a RADIUS client	49
Displaying and maintaining AAA	50
AAA configuration examples	50
AAA for Telnet users by an HWTACACS server	50
AAA for Telnet users by separate servers	51
Authentication/authorization for SSH/Telnet users by a RADIUS server	53
AAA for portal users by a RADIUS server	56
AAA for 802.1X users by a RADIUS server	65
Level switching authentication for Telnet users by an HWTACACS server	71
RADIUS authentication and authorization for Telnet users by a switch	74
Troubleshooting AAA	76
Troubleshooting RADIUS	76
Troubleshooting HWTACACS	77
802.1X fundamentals	78
802.1X architecture	78
Controlled/uncontrolled port and port authorization status	78
802.1X-related protocols	79
Packet formats	80
EAP over RADIUS	81
Initiating 802.1X authentication	81
802.1X client as the initiator	81

Access device as the initiator	82
802.1X authentication procedures	82
A comparison of EAP relay and EAP termination	83
EAP relay	83
EAP termination	86
Configuring 802.1X	87
HP implementation of 802.1X	87
Access control methods	87
Using 802.1X authentication with other features	87
Configuration prerequisites	92
802.1X configuration task list	92
Enabling 802.1X	93
Configuration guidelines	93
Configuration procedure	93
Enabling EAP relay or EAP termination	93
Setting the port authorization state	94
Specifying an access control method	95
Setting the maximum number of concurrent 802.1X users on a port	95
Setting the maximum number of authentication request attempts	96
Setting the 802.1X authentication timeout timers	96
Configuring the online user handshake function	96
Configuration guidelines	97
Configuration procedure	97
Configuring the authentication trigger function	97
Configuration guidelines	98
Configuration procedure	98
Specifying a mandatory authentication domain on a port	98
Configuring the quiet timer	99
Enabling the periodic online user re-authentication function	99
Configuration guidelines	99
Configuration procedure	99
Configuring an 802.1X guest VLAN	100
Configuration guidelines	100
Configuration prerequisites	100
Configuration procedure	101
Configuring an Auth-Fail VLAN	101
Configuration guidelines	101
Configuration prerequisites	102
Configuration procedure	102
Configuring an 802.1X critical VLAN	102
Configuration guidelines	102
Configuration prerequisites	102
Configuration procedure	102
Specifying supported domain name delimiters	103
Displaying and maintaining 802.1X	103
802.1X authentication configuration example	104
Network requirements	104
Configuration procedure	104
Verifying the configuration	106
802.1X with guest VLAN and VLAN assignment configuration example	106
Network requirements	106
Configuration procedure	107
Verifying the configuration	108
802.1X with ACL assignment configuration example	109

Network requirements	109
Configuration procedure	109
Verifying the configuration	110
Configuring EAD fast deployment	111
Overview	111
Free IP	111
URL redirection	111
Configuration prerequisites	111
Configuring a free IP	111
Configuring the redirect URL	112
Setting the EAD rule timer	112
Displaying and maintaining EAD fast deployment	112
EAD fast deployment configuration example	113
Network requirements	113
Configuration procedure	114
Verifying the configuration	114
Troubleshooting EAD fast deployment	115
Web browser users cannot be correctly redirected	115
Configuring MAC authentication	116
MAC authentication overview	116
User account policies	116
Authentication approaches	116
MAC authentication timers	117
Using MAC authentication with other features	117
VLAN assignment	117
ACL assignment	117
Guest VLAN	117
Critical VLAN	118
Configuration task list	118
Basic configuration for MAC authentication	118
Specifying a MAC authentication domain	120
Configuring a MAC authentication guest VLAN	120
Configuring a MAC authentication critical VLAN	121
Displaying and maintaining MAC authentication	122
MAC authentication configuration examples	122
Local MAC authentication configuration example	122
RADIUS-based MAC authentication configuration example	124
ACL assignment configuration example	126
Configuring portal authentication	129
Overview	129
Extended portal functions	129
Portal system components	129
Portal system using the local portal server	131
Portal authentication modes	132
Portal support for EAP (available only on the HP 5500 EI series)	133
Layer 2 portal authentication process	134
Layer 3 portal authentication process (available only on the HP 5500 EI series)	135
Portal stateful failover (available only on the HP 5500 EI series)	138
Portal authentication across VPNs (available only on the HP 5500 EI series)	140
Portal configuration task list	140
Configuration prerequisites	141
Specifying the portal server	142
Specifying the local portal server for Layer 2 portal authentication	142

Specifying a portal server for Layer 3 portal authentication (available only on the HP 5500 EI series) ..	143
Configuring the local portal server ..	143
Customizing authentication pages ..	143
Configuring the local portal server ..	146
Enabling portal authentication ..	147
Enabling Layer 2 portal authentication ..	147
Enabling Layer 3 portal authentication (available only on the HP 5500 EI series) ..	148
Controlling access of portal users ..	149
Configuring a portal-free rule ..	149
Configuring an authentication source subnet (available only on the HP 5500 EI series) ..	150
Setting the maximum number of online portal users ..	150
Specifying an authentication domain for portal users ..	151
Configuring Layer 2 portal authentication to support web proxy ..	151
Enabling support for portal user moving ..	152
Specifying an Auth-Fail VLAN for portal authentication ..	152
Configuring RADIUS related attributes ..	153
Specifying NAS-Port-Type for an interface ..	153
Specifying a NAS ID profile for an interface ..	154
Specifying a source IP address for outgoing portal packets ..	154
Configuring portal stateful failover (available only on the HP 5500 EI series) ..	155
Specifying an auto redirection URL for authenticated portal users ..	156
Configuring portal detection functions ..	157
Configuring online Layer 2 portal user detection ..	157
Configuring the portal server detection function (available only on the HP 5500 EI series) ..	157
Configuring portal user information synchronization (available only on the HP 5500 EI series) ..	159
Logging off portal users ..	160
Displaying and maintaining portal ..	160
Portal configuration examples ..	161
Configuring direct portal authentication ..	161
Configuring re-DHCP portal authentication ..	166
Configuring cross-subnet portal authentication ..	168
Configuring direct portal authentication with extended functions ..	170
Configuring re-DHCP portal authentication with extended functions ..	172
Configuring cross-subnet portal authentication with extended functions ..	174
Configuring portal stateful failover ..	176
Configuring portal server detection and portal user information synchronization ..	184
Configuring Layer 2 portal authentication ..	189
Troubleshooting portal ..	193
Inconsistent keys on the access device and the portal server ..	193
Incorrect server port number on the access device ..	193
Configuring triple authentication ..	195
Overview ..	195
Triple authentication mechanism ..	195
Using triple authentication with other features ..	196
Configuring triple authentication ..	196
Triple authentication configuration examples ..	197
Triple authentication basic function configuration example ..	197
Triple authentication supporting VLAN assignment and Auth-Fail VLAN configuration example ..	199
Configuring port security ..	205
Overview ..	205
Port security features ..	205
Port security modes ..	205
Working with guest VLAN and Auth-Fail VLAN ..	208

Configuration task list	208
Enabling port security	209
Setting port security's limit on the number of MAC addresses on a port	209
Setting the port security mode	210
Configuration prerequisites	210
Configuration procedure	210
Configuring port security features	211
Configuring NTK	211
Configuring intrusion protection	211
Enabling port security traps	212
Configuring secure MAC addresses	212
Configuration prerequisites	213
Configuration procedure	213
Ignoring authorization information from the server	214
Displaying and maintaining port security	214
Port security configuration examples	215
Configuring the autoLearn mode	215
Configuring the userLoginWithOUI mode	217
Configuring the macAddressElseUserLoginSecure mode	222
Troubleshooting port security	224
Cannot set the port security mode	224
Cannot configure secure MAC addresses	225
Cannot change port security mode when a user is online	225
Configuring a user profile	227
User profile overview	227
User profile configuration task list	227
Creating a user profile	227
Configuration prerequisites	227
Configuration procedure	227
Applying a QoS policy	228
Configuration guidelines	228
Configuration procedure	228
Enabling a user profile	228
Displaying and maintaining user profiles	229
Configuring password control	230
Password control overview	230
Password control configuration task list	232
Configuring password control	233
Enabling password control	233
Setting global password control parameters	233
Setting user group password control parameters	235
Setting local user password control parameters	235
Setting super password control parameters	236
Setting a local user password in interactive mode	236
Displaying and maintaining password control	236
Password control configuration example	237
Configuring HABP	240
HABP overview	240
Configuring HABP	241
Configuring the HABP server	241
Configuring an HABP client	241
Displaying and maintaining HABP	242
HABP configuration example	242

Managing public keys	245
Overview	245
Configuration task list	245
Creating a local asymmetric key pair	246
Displaying or exporting the local host public key	246
Destroying a local asymmetric key pair	248
Specifying the peer public key on the local device	248
Displaying and maintaining public keys	249
Public key configuration examples	249
Manually specifying the peer public key on the local device	249
Importing a peer public key from a public key file	251
Configuring PKI	254
Overview	254
PKI terms	254
PKI architecture	255
PKI applications	255
How PKI operates	256
PKI configuration task list	256
Configuring an entity DN	257
Configuring a PKI domain	258
Configuration guidelines	258
Configuration procedure	259
Submitting a PKI certificate request	259
Submitting a certificate request in auto mode	260
Submitting a certificate request in manual mode	260
Retrieving a certificate manually	261
Configuration guidelines	261
Configuration procedure	262
Configuring PKI certificate verification	262
Configuration guidelines	262
Configuring CRL-checking-enabled PKI certificate verification	262
Configuring CRL-checking-disabled PKI certificate verification	263
Destroying a local RSA key pair	263
Deleting a certificate	263
Configuring an access control policy	264
Displaying and maintaining PKI	264
PKI configuration examples	265
Requesting a certificate from a CA server running RSA Keon	265
Requesting a certificate from a CA server running Windows 2003 Server	268
Configuring a certificate attribute-based access control policy	271
Troubleshooting PKI	273
Failed to retrieve a CA certificate	273
Failed to request a local certificate	273
Failed to retrieve CRLs	274
Configuring IPsec	275
Overview	275
IPsec implementation	275
Basic concepts	276
IPsec for IPv6 routing protocols	278
Protocols and standards	278
Configuring IPsec for IPv6 routing protocols	278
Configuring an IPsec proposal	278
Configuring an IPsec policy	280

Displaying and maintaining IPsec	281
IPsec for RIPng configuration example	281
Configuring SSH2.0	286
Overview	286
Introduction to SSH2.0	286
SSH operation	286
SSH connection across VPNs (available only on the HP 5500 EI)	289
Configuring the switch as an SSH server	289
SSH server configuration task list	289
Generating DSA or RSA key pairs	289
Enabling the SSH server function	290
Configuring the user interfaces for SSH clients	290
Configuring a client public key	291
Configuring an SSH user	292
Setting the SSH management parameters	293
Setting the DSCP value for packets sent by the SSH server	294
Configuring the switch as an SSH client	294
SSH client configuration task list	294
Specifying a source IP address/interface for the SSH client	294
Configuring whether first-time authentication is supported	295
Establishing a connection between the SSH client and server	296
Setting the DSCP value for packets sent by the SSH client	296
Displaying and maintaining SSH	296
SSH server configuration examples	297
When the switch acts as a server for password authentication	297
When the switch acts as a server for publickey authentication	299
SSH client configuration examples	304
When switch acts as client for password authentication	304
When switch acts as client for publickey authentication	307
Configuring SFTP	310
Overview	310
Configuring the switch as an SFTP server	310
Enabling the SFTP server	310
Configuring the SFTP connection idle timeout period	310
Configuring the switch as an SFTP client	311
Specifying a source IP address or interface for the SFTP client	311
Establishing a connection to the SFTP server	311
Working with SFTP directories	312
Working with SFTP files	313
Displaying help information	313
Terminating the connection to the remote SFTP server	314
Setting the DSCP value for packets sent by the SFTP client	314
SFTP client configuration example	314
SFTP server configuration example	318
Configuring SCP	321
Overview	321
Configuring the switch as an SCP server	321
Configuring the switch as the SCP client	321
SCP client configuration example	322
SCP server configuration example	323
Configuring SSL	325
Overview	325

SSL security mechanism	325
SSL protocol stack	325
Configuration task list	326
Configuring an SSL server policy	326
SSL server policy configuration example	327
Configuring an SSL client policy	329
Displaying and maintaining SSL	330
Troubleshooting SSL	330
Configuring TCP attack protection	332
Overview	332
Enabling the SYN Cookie feature	332
Displaying and maintaining TCP attack protection	332
Configuring IP source guard	334
Overview	334
Static IP source guard entries	334
Dynamic IP source guard entries	335
Configuration task list	335
Configuring the IPv4 source guard function	336
Configuring IPv4 source guard on a port	336
Configuring a static IPv4 source guard entry	337
Setting the maximum number of IPv4 source guard entries	338
Configuring the IPv6 source guard function	338
Configuring IPv6 source guard on a port	338
Configuring a static IPv6 source guard entry	339
Setting the maximum number of IPv6 source guard entries	340
Displaying and maintaining IP source guard	341
IP source guard configuration examples	341
Static IPv4 source guard configuration example	341
Dynamic IPv4 source guard using DHCP snooping configuration example	343
Dynamic IPv4 source guard using DHCP relay configuration example	345
Static IPv6 source guard configuration example	346
Dynamic IPv6 source guard using DHCPv6 snooping configuration example	346
Dynamic IPv6 source guard using ND snooping configuration example	348
Global static IP source guard configuration example	349
Troubleshooting IP source guard	350
Configuring ARP attack protection	351
Overview	351
ARP attack protection configuration task list	351
Configuring ARP defense against IP packet attacks	352
Configuring ARP source suppression	353
Enabling ARP black hole routing	353
Displaying and maintaining ARP defense against IP packet attacks	353
Configuration example	353
Configuring ARP packet rate limit	355
Introduction	355
Configuration procedure	355
Configuring source MAC address based ARP attack detection	356
Configuration procedure	356
Displaying and maintaining source MAC address based ARP attack detection	356
Configuration example	357
Configuring ARP packet source MAC address consistency check	358
Introduction	358
Configuration procedure	358

Configuring ARP active acknowledgement	358
Introduction	358
Configuration procedure	358
Configuring ARP detection	359
Introduction	359
Configuring user validity check	359
Configuring ARP packet validity check	360
Configuring ARP restricted forwarding	361
Displaying and maintaining ARP detection	361
User validity check configuration example	362
User validity check and ARP packet validity check configuration example	363
ARP restricted forwarding configuration example	364
Configuring ARP automatic scanning and fixed ARP	366
Configuration guidelines	366
Configuration procedure	367
Configuring ARP gateway protection	367
Configuration guidelines	367
Configuration procedure	367
Configuration example	368
Configuring ARP filtering	368
Configuration guidelines	369
Configuration procedure	369
Configuration example	369
Configuring ND attack defense	371
Overview	371
Enabling source MAC consistency check for ND packets	372
Configuring the ND detection function	372
Introduction to ND detection	372
Configuration guidelines	373
Configuration procedure	373
Displaying and maintaining ND detection	373
ND detection configuration example	374
Network requirements	374
Configuration procedure	374
Configuring URPF (available only on the HP 5500 EI)	376
URPF overview	376
What is URPF	376
URPF check modes	376
How URPF works	377
Network application	379
Configuring URPF	379
URPF configuration example	379
Configuring SAVI	381
SAVI overview	381
Configuring global SAVI	381
SAVI configuration in DHCPv6-only address assignment scenario	382
SAVI configuration in SLAAC-only address assignment scenario	384
SAVI configuration in DHCPv6+SLAAC address assignment scenario	386
Configuring blacklist	388
Overview	388
Configuring the blacklist feature	388
Displaying and maintaining the blacklist	388

Blacklist configuration example	389
Network requirements	389
Configuration procedure	389
Verifying the configuration	389

Index	391
-------------	-----

Configuring AAA

In the HP 5500 Switch Series, only the HP 5500 EI switches support MCE and VPN configurations.

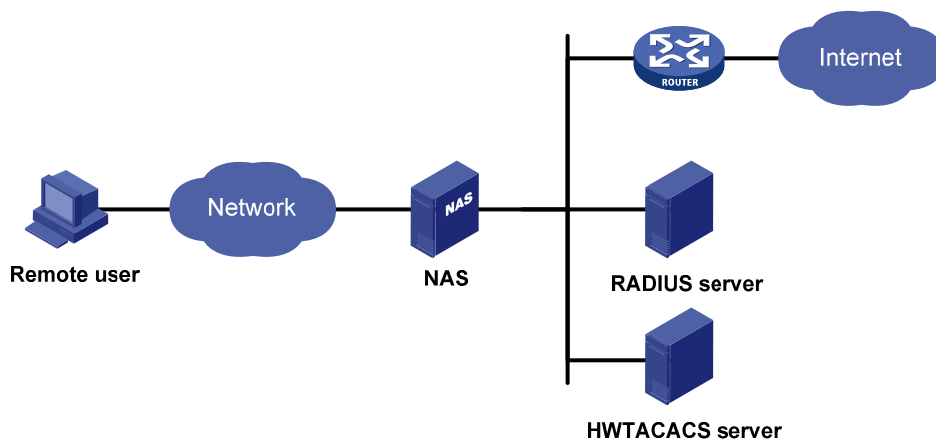
AAA overview

Authentication, Authorization, and Accounting (AAA) provides a uniform framework for implementing network access management. It can provide the following security functions:

- **Authentication**—Identifies users and determines whether a user is valid.
- **Authorization**—Grants different users different rights and controls their access to resources and services. For example, a user who has successfully logged in to the switch can be granted read and print permissions to the files on the switch.
- **Accounting**—Records all user network service usage information, including the service type, start time, and traffic. The accounting function not only provides the information required for charging, but also allows for network security surveillance.

AAA usually uses a client/server model. The client runs on the network access server (NAS), which is also referred to as the access device. The server maintains user information centrally. In an AAA network, a NAS is a server for users but a client for the AAA servers. See [Figure 1](#).

Figure 1 Network diagram



When a user tries to log in to the NAS, use network resources, or access other networks, the NAS authenticates the user. The NAS can transparently pass the user's authentication, authorization, and accounting information to the servers. The RADIUS and HWTACACS protocols define how a NAS and a remote server exchange user information between them.

In the network shown in [Figure 1](#), there is a RADIUS server and an HWTACACS server. You can choose different servers for different security functions. For example, you can use the HWTACACS server for authentication and authorization, and the RADIUS server for accounting.

You can choose the three security functions provided by AAA as needed. For example, if your company only wants employees to be authenticated before they access specific resources, configure an authentication server. If network usage information is needed, you must also configure an accounting server.

AAA can be implemented through multiple protocols. The switch supports using RADIUS and HWTACACS. RADIUS is often used in practice.

RADIUS

Remote Authentication Dial-In User Service (RADIUS) is a distributed information interaction protocol that uses a client/server model. It can protect networks against unauthorized access and is often used in network environments where both high security and remote user access are required.

RADIUS uses UDP as the transport protocol. It uses UDP port 1812 for authentication and UDP port 1813 for accounting.

RADIUS was originally designed for dial-in user access. With the addition of new access methods, RADIUS has been extended to support additional access methods, such as Ethernet and ADSL. RADIUS provides access authentication and authorization services, and its accounting function collects and records network resource usage information.

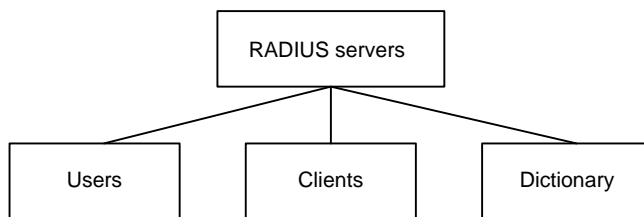
Client/server model

The RADIUS client runs on the NASs located throughout the network. It passes user information to designated RADIUS servers and acts on the responses (for example, rejects or accepts user access requests).

The RADIUS server runs on the computer or workstation at the network center and maintains information related to user authentication and network service access. It listens to connection requests, authenticates users, and returns user access control information (for example, rejecting or accepting the user access request) to the clients.

In general, the RADIUS server maintains the following databases: Users, Clients, and Dictionary.

Figure 2 RADIUS server components



- **Users**—Stores user information, such as usernames, passwords, applied protocols, and IP addresses.
- **Clients**—Stores information about RADIUS clients, such as shared keys and IP addresses.
- **Dictionary**—Stores RADIUS protocol attributes and their values.

Security and authentication mechanisms

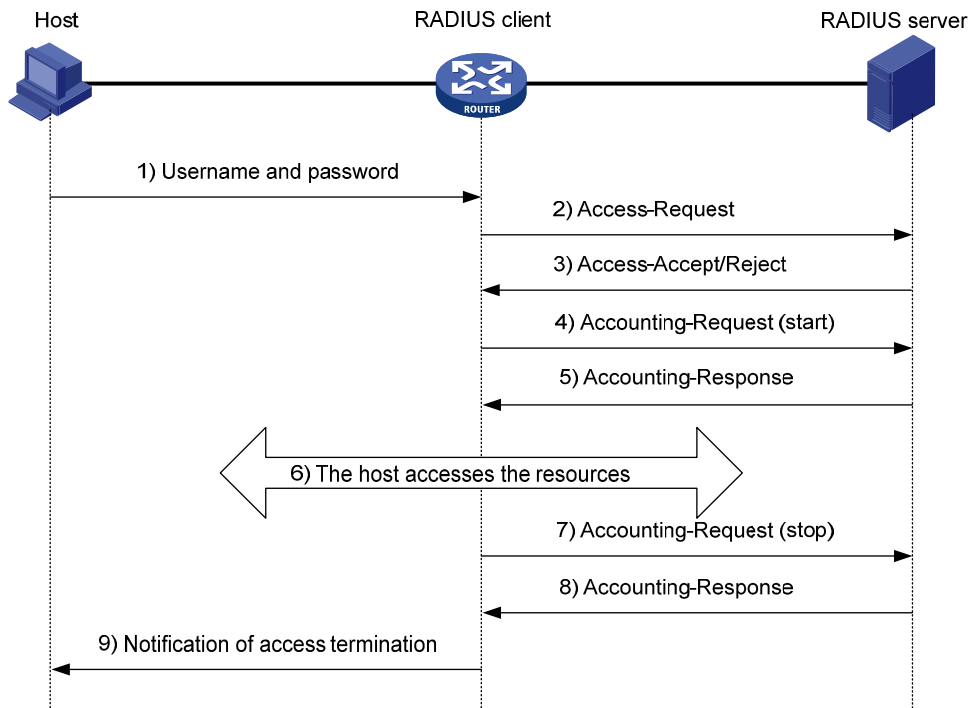
A RADIUS client and the RADIUS server use the shared key to authenticate RADIUS packets and encrypt user passwords that are exchanged between them. The keys are never transmitted over the network. This security mechanism improves the security of RADIUS communication and prevents user passwords from being intercepted on insecure networks.

A RADIUS server supports multiple user authentication methods. A RADIUS server can also act as the client of another AAA server to provide authentication proxy services.

Basic RADIUS message exchange process

Figure 3 illustrates the interactions between the host, the RADIUS client, and the RADIUS server.

Figure 3 Basic RADIUS message exchange process



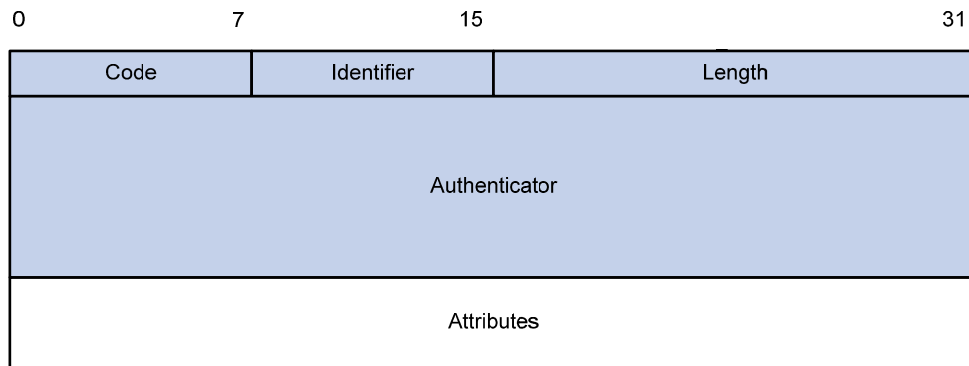
RADIUS operates in the following manner:

1. The host initiates a connection request that carries the user's username and password to the RADIUS client.
2. Having received the username and password, the RADIUS client sends an authentication request (Access-Request) to the RADIUS server, with the user password encrypted by using the Message-Digest 5 (MD5) algorithm and the shared key.
3. The RADIUS server authenticates the username and password. If the authentication succeeds, the server sends back an Access-Accept message containing the user's authorization information. If the authentication fails, the server returns an Access-Reject message.
4. The RADIUS client permits or denies the user according to the returned authentication result. If it permits the user, it sends a start-accounting request (Accounting-Request) to the RADIUS server.
5. The RADIUS server returns a start-accounting response (Accounting-Response) and starts accounting.
6. The user accesses the network resources.
7. The host requests the RADIUS client to tear down the connection and the RADIUS client sends a stop-accounting request (Accounting-Request) to the RADIUS server.
8. The RADIUS server returns a stop-accounting response (Accounting-Response) and stops accounting for the user.

RADIUS packet format

RADIUS uses UDP to transmit messages. To ensure smooth message exchange between the RADIUS server and the client, RADIUS uses a series of mechanisms, including the timer management mechanism, the retransmission mechanism, and the backup server mechanism. Figure 4 shows the RADIUS packet format.

Figure 4 RADIUS packet format



Descriptions of the fields are as follows:

- The Code field (1 byte long) indicates the type of the RADIUS packet. [Table 1](#) gives the possible values and their meanings.

Table 1 Main values of the Code field

Code	Packet type	Description
1	Access-Request	From the client to the server. A packet of this type carries user information for the server to authenticate the user. It must contain the User-Name attribute and can optionally contain the attributes of NAS-IP-Address, User-Password, and NAS-Port.
2	Access-Accept	From the server to the client. If all the attribute values carried in the Access-Request are acceptable, the authentication succeeds, and the server sends an Access-Accept response.
3	Access-Reject	From the server to the client. If any attribute value carried in the Access-Request is unacceptable, the authentication fails and the server sends an Access-Reject response.
4	Accounting-Request	From the client to the server. A packet of this type carries user information for the server to start or stop accounting for the user. The Acct-Status-Type attribute in the packet indicates whether to start or stop accounting.
5	Accounting-Response	From the server to the client. The server sends a packet of this type to notify the client that it has received the Accounting-Request and has successfully recorded the accounting information.

- The Identifier field (1 byte long) is used to match request and response packets and to detect duplicate request packets. Request and response packets of the same type have the same identifier.
- The Length field (2 bytes long) indicates the length of the entire packet, including the Code, Identifier, Length, Authenticator, and Attribute fields. Bytes beyond this length are considered padding and are ignored at the receiver. If the length of a received packet is less than this length, the packet is dropped. The value of this field is in the range of 20 to 4096.
- The Authenticator field (16 bytes long) is used to authenticate replies from the RADIUS server and to encrypt user passwords. There are two types of authenticators: request authenticator and response authenticator.

- The Attributes field (variable in length) carries the specific authentication, authorization, and accounting information that defines the configuration details of the request or response. This field may contain multiple attributes, each with three sub-fields:
 - **Type**—(1 byte long) Type of the attribute. It is in the range of 1 to 255. Commonly used RADIUS attributes are defined in RFC 2865, RFC 2866, RFC 2867, and RFC 2868. [Table 2](#) shows a list of the attributes. For more information about commonly used standard RADIUS attributes, see "[Commonly used standard RADIUS attributes.](#)"
 - **Length**—(1 byte long) Length of the attribute in bytes, including the Type, Length, and Value fields.
 - **Value**—(Up to 253 bytes) Value of the attribute. Its format and content depend on the Type and Length fields.

Table 2 Commonly used RADIUS attributes

No.	Attribute	No.	Attribute
1	User-Name	45	Acct-Authentic
2	User-Password	46	Acct-Session-Time
3	CHAP-Password	47	Acct-Input-Packets
4	NAS-IP-Address	48	Acct-Output-Packets
5	NAS-Port	49	Acct-Terminate-Cause
6	Service-Type	50	Acct-Multi-Session-Id
7	Framed-Protocol	51	Acct-Link-Count
8	Framed-IP-Address	52	Acct-Input-Gigawords
9	Framed-IP-Netmask	53	Acct-Output-Gigawords
10	Framed-Routing	54	(unassigned)
11	Filter-ID	55	Event-Timestamp
12	Framed-MTU	56-59	(unassigned)
13	Framed-Compression	60	CHAP-Challenge
14	Login-IP-Host	61	NAS-Port-Type
15	Login-Service	62	Port-Limit
16	Login-TCP-Port	63	Login-LAT-Port
17	(unassigned)	64	Tunnel-Type
18	Reply-Message	65	Tunnel-Medium-Type
19	Callback-Number	66	Tunnel-Client-Endpoint
20	Callback-ID	67	Tunnel-Server-Endpoint
21	(unassigned)	68	Acct-Tunnel-Connection
22	Framed-Route	69	Tunnel-Password
23	Framed-IPX-Network	70	ARAP-Password
24	State	71	ARAP-Features
25	Class	72	ARAP-Zone-Access
26	Vendor-Specific	73	ARAP-Security

No.	Attribute	No.	Attribute
27	Session-Timeout	74	ARAP-Security-Data
28	Idle-Timeout	75	Password-Retry
29	Termination-Action	76	Prompt
30	Called-Station-Id	77	Connect-Info
31	Calling-Station-Id	78	Configuration-Token
32	NAS-Identifier	79	EAP-Message
33	Proxy-State	80	Message-Authenticator
34	Login-LAT-Service	81	Tunnel-Private-Group-id
35	Login-LAT-Node	82	Tunnel-Assignment-id
36	Login-LAT-Group	83	Tunnel-Preference
37	Framed-AppleTalk-Link	84	ARAP-Challenge-Response
38	Framed-AppleTalk-Network	85	Acct-Interim-Interval
39	Framed-AppleTalk-Zone	86	Acct-Tunnel-Packets-Lost
40	Acct-Status-Type	87	NAS-Port-Id
41	Acct-Delay-Time	88	Framed-Pool
42	Acct-Input-Octets	89	(unassigned)
43	Acct-Output-Octets	90	Tunnel-Client-Auth-id
44	Acct-Session-Id	91	Tunnel-Server-Auth-id

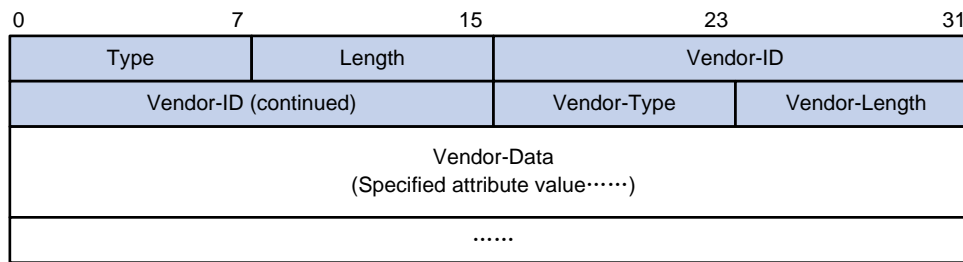
Extended RADIUS attributes

The RADIUS protocol features excellent extensibility. Attribute 26 (Vendor-Specific), an attribute defined by RFC 2865, allows a vendor to define extended attributes to implement functions that the standard RADIUS protocol does not provide.

A vendor can encapsulate multiple sub-attributes in the type-length-value (TLV) format in RADIUS packets for extension of applications. As shown in [Figure 5](#), a sub-attribute encapsulated in Attribute 26 consists of the following parts:

- **Vendor-ID**—Indicates the ID of the vendor. Its most significant byte is 0, and the other three bytes contains a code that is compliant to RFC 1700. For more information about the proprietary RADIUS sub-attributes of HP, see "[HP proprietary RADIUS sub-attributes](#)."
- **Vendor-Type**—Indicates the type of the sub-attribute.
- **Vendor-Length**—Indicates the length of the sub-attribute.
- **Vendor-Data**—Indicates the contents of the sub-attribute.

Figure 5 Segment of a RADIUS packet containing an extended attribute



HWTACACS

HW Terminal Access Controller Access Control System (HWTACACS) is an enhanced security protocol based on TACACS (RFC 1492). Similar to RADIUS, it uses a client/server model for information exchange between the NAS and the HWTACACS server.

HWTACACS typically provides AAA services for Point-to-Point Protocol (PPP) users, Virtual Private Dial-up Network (VPDN) users, and terminal users. In a typical HWTACACS scenario, some terminal users log in to the NAS for operations. Working as the HWTACACS client, the NAS sends the usernames and passwords of the users to the HWTACACS sever for authentication. After passing authentication and being authorized, the users log in to the switch and performs operations, and the HWTACACS server records the operations that each user performs.

Differences between HWTACACS and RADIUS

HWTACACS and RADIUS both provide authentication, authorization, and accounting services. They have many features in common, such as using a client/server model, using shared keys for user information security, and providing flexibility and extensibility.

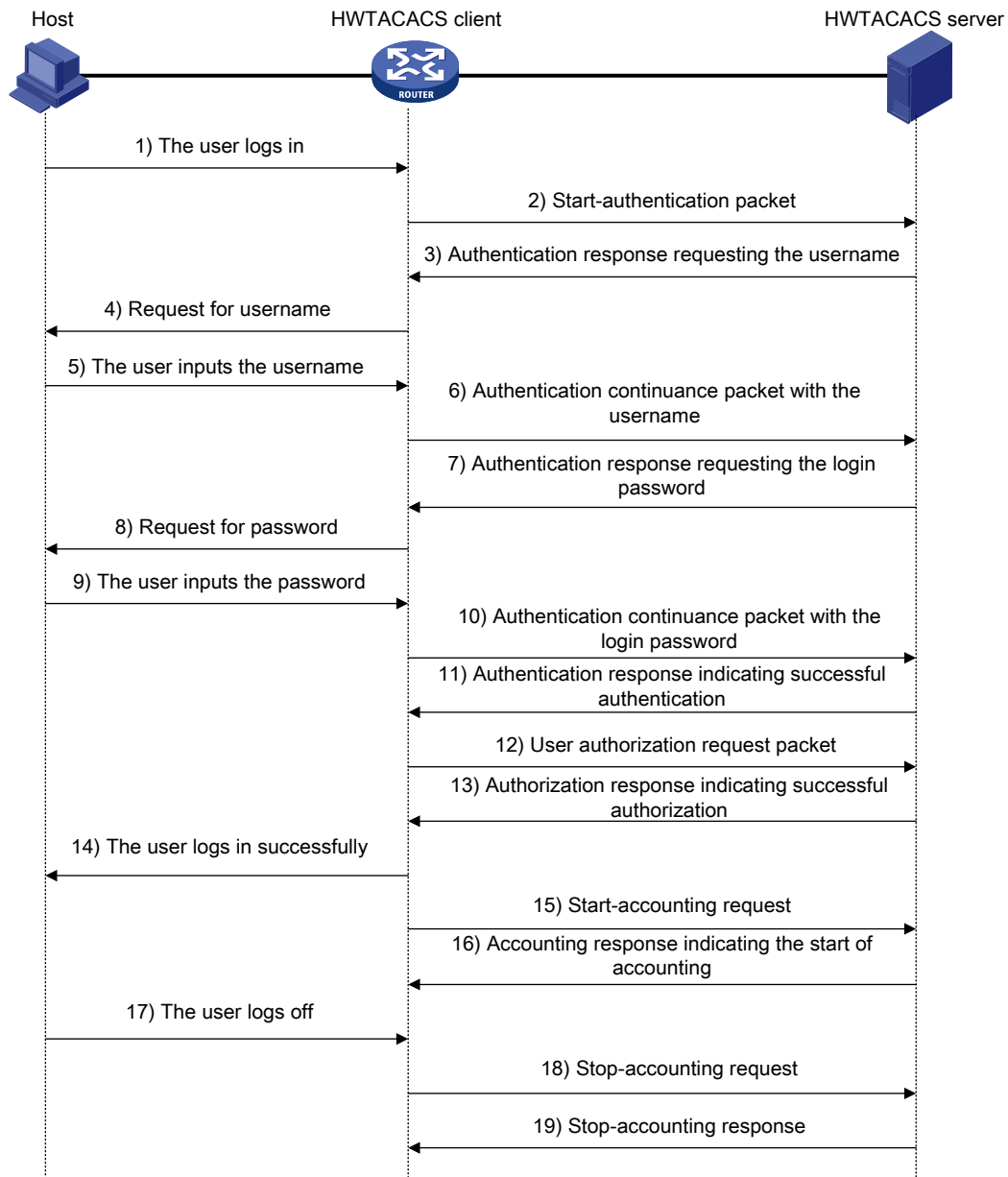
Table 3 Primary differences between HWTACACS and RADIUS

HWTACACS	RADIUS
Uses TCP, providing more reliable network transmission.	Uses UDP, providing higher transport efficiency.
Encrypts the entire packet except for the HWTACACS header.	Encrypts only the user password field in an authentication packet.
Protocol packets are complicated and authorization is independent of authentication. Authentication and authorization can be deployed on different HWTACACS servers.	Protocol packets are simple and the authorization process is combined with the authentication process.
Supports authorization of configuration commands. Which commands a user can use depends on both the user level and the AAA authorization. A user can use only commands that are at, or lower than, the user level and authorized by the HWTACACS server.	Does not support authorization of configuration commands. Which commands a user can use solely depends on the level of the user. A user can use all the commands at, or lower than, the user level.

Basic HWTACACS message exchange process

The following example describes how HWTACACS performs user authentication, authorization, and accounting for a Telnet user.

Figure 6 Basic HWTACACS message exchange process for a Telnet user



HWTACACS operates in the following manner:

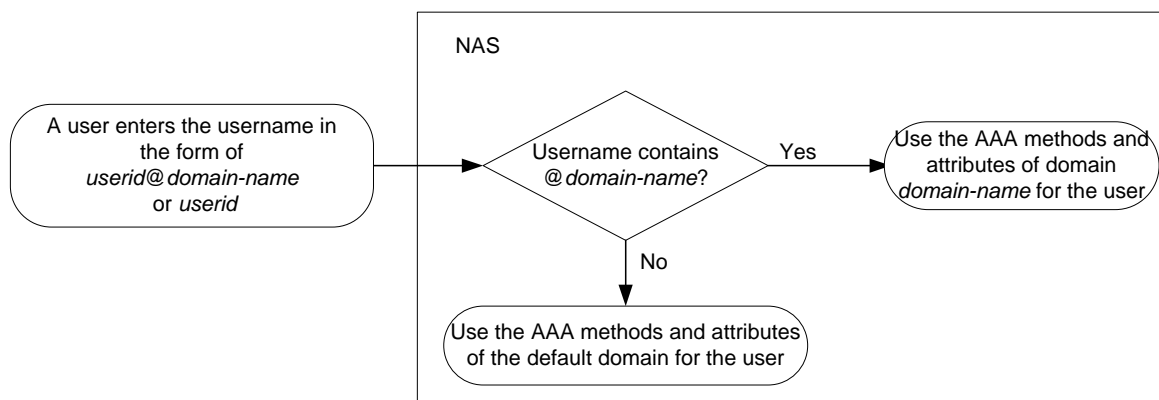
1. A Telnet user sends an access request to the HWTACACS client.
2. Upon receiving the request, the HWTACACS client sends a start-authentication packet to the HWTACACS server.
3. The HWTACACS server sends back an authentication response to request the username.
4. Upon receiving the response, the HWTACACS client asks the user for the username.
5. The user enters the username.
6. After receiving the username from the user, the HWTACACS client sends the server a continue-authentication packet that carries the username.
7. The HWTACACS server sends back an authentication response, requesting the login password.
8. Upon receipt of the response, the HWTACACS client asks the user for the login password.

9. The user enters the password.
10. After receiving the login password, the HWTACACS client sends the HWTACACS server a continue-authentication packet that carries the login password.
11. The HWTACACS server sends back an authentication response to indicate that the user has passed authentication.
12. The HWTACACS client sends the user authorization request packet to the HWTACACS server.
13. The HWTACACS server sends back the authorization response, indicating that the user is now authorized.
14. Knowing that the user is now authorized, the HWTACACS client pushes its configuration interface to the user.
15. The HWTACACS client sends a start-accounting request to the HWTACACS server.
16. The HWTACACS server sends back an accounting response, indicating that it has received the start-accounting request.
17. The user logs off.
18. The HWTACACS client sends a stop-accounting request to the HWTACACS server.
19. The HWTACACS server sends back a stop-accounting response, indicating that the stop-accounting request has been received.

Domain-based user management

A NAS manages users based on Internet service provider (ISP) domains. On a NAS, each user belongs to one ISP domain. A NAS determines the ISP domain a user belongs to by the username entered by the user at login, as shown in [Figure 7](#).

Figure 7 Determining the ISP domain of a user by the username



The authentication, authorization, and accounting of a user depends on the AAA methods configured for the domain to which the user belongs. If no specific AAA methods are configured for the domain, the default methods are used. By default, a domain uses local authentication, local authorization, and local accounting.

AAA allows you to manage users based on their access types:

- **LAN users**—Users on a LAN who must pass 802.1X or MAC address authentication to access the network.
- **Login users**—Users who want to log in to the switch, including SSH users, Telnet users, Web users, FTP users, and terminal users.

- **Portal users**—Users who must pass portal authentication to access the network.

In addition, AAA provides the following services for login users to enhance switch security:

- **Command authorization**—Enables the NAS to defer to the authorization server to determine whether a command entered by a login user is permitted for the user, making sure that login users execute only commands they are authorized to execute. For more information about command authorization, see *Fundamentals Configuration Guide*.
- **Command accounting**—Allows the accounting server to record all commands executed on the switch or all authorized commands successfully executed. For more information about command accounting, see *Fundamentals Configuration Guide*.
- **Level switching authentication**—Allows the authentication server to authenticate users who perform privilege level switching. As long as passing level switching authentication, users can switch their user privilege levels, without logging out and disconnecting current connections. For more information about user privilege level switching, see *Fundamentals Configuration Guide*.

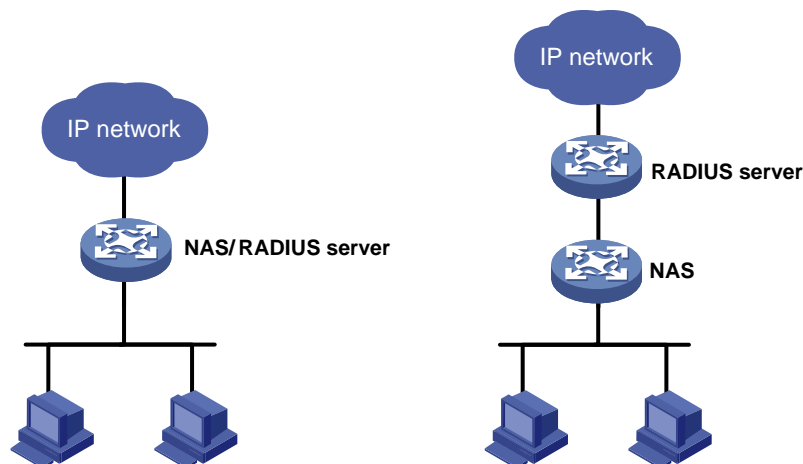
You can configure different authentication, authorization, and accounting methods for different types of users in a domain. See "[Configuring AAA methods for ISP domains](#)."

RADIUS server feature of the switch

Generally, the RADIUS server runs on a computer or workstation, and the RADIUS client runs on a NAS. A network device that supports the RADIUS server feature can also serve as the RADIUS server, working with RADIUS clients to implement user authentication, authorization, and accounting. As shown in [Figure 8](#), the RADIUS server and client can reside on the same switch or different switches.

Using a network device as the RADIUS server simplifies networking and reduces deployment costs. This implementation is usually deployed on networks by using the clustering feature. In such a scenario, configure the RADIUS server feature on a management device at the distribution layer, so that the device functions as a RADIUS server to cooperate with cluster member switches at the access layer to provide user authentication and authorization services.

Figure 8 Devices functioning as a RADIUS server



The switch can serve as a RADIUS server to provide the following functions:

- User information management
You can create, modify, and delete user information, including the username, password, authority, lifetime, and user description.
- RADIUS client information management

You can create and delete RADIUS clients, which are identified by IP addresses and configured with attributes such as a shared key. With a managed client range configured, the RADIUS server processes only the RADIUS packets from the clients within the management range. A shared key is used to ensure secure communication between a RADIUS client and the RADIUS server.

- RADIUS authentication and authorization

With the RADIUS server enabled, the switch checks whether or not the client of an incoming RADIUS packet is under its management. If yes, it verifies the packet validity by using the shared key, checks whether there is an account with the username, whether the password is correct, and whether the user attributes meet the requirements defined on the RADIUS server (for example, whether the account has expired). Then, the RADIUS server assigns the corresponding authority to the client if the authentication succeeds, or denies the client if the authentication fails.

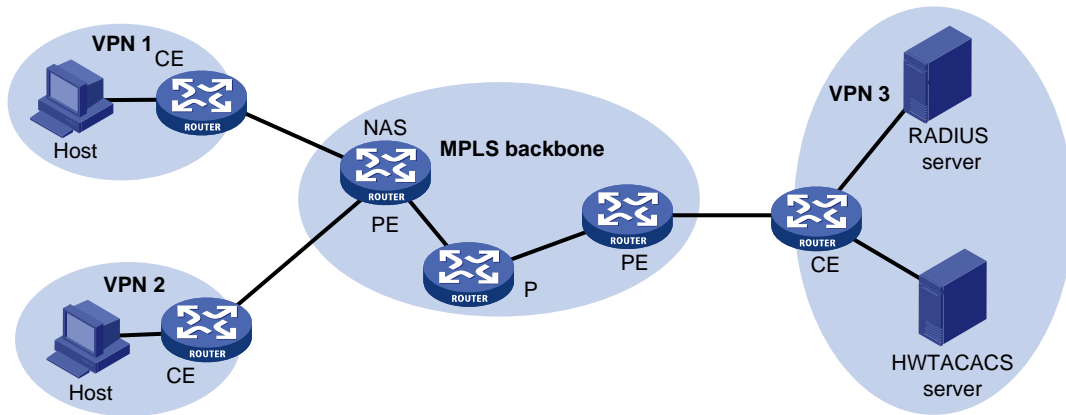
NOTE:

A RADIUS server running the standard RADIUS protocol listens on UDP port 1812 for authentication requests, but an HP switch listens on UDP port 1645 instead when acting as the RADIUS server. Be sure to specify 1645 as the authentication port number on the RADIUS client when you use an HP switch as the RADIUS server.

AAA for MPLS L3VPNs (available only on the HP 5500 EI)

In an MPLS L3VPN scenario where clients in different VPNs are centrally authenticated, you can deploy AAA across VPNs to enable forwarding RADIUS and HWTACACS packets across MPLS VPNs. With the AAA across VPNs feature, the PE at the left side of the MPLS backbone serves as a NAS and transparently delivers the AAA packets of private users in VPN 1 and VPN 2 to the AAA servers in VPN 3 for centralized authentication, as shown in [Figure 9](#). Authentication packets of private users in different VPNs do not affect each other.

Figure 9 Network diagram



NOTE:

This feature can also help an MCE to implement portal authentication for VPNs. For more information about MCE, see *Layer 3 - IP Routing Configuration Guide*.

Protocols and standards

The following protocols and standards are related to AAA, RADIUS, and HWTACACS:

- RFC 2865, *Remote Authentication Dial In User Service (RADIUS)*
- RFC 2866, *RADIUS Accounting*
- RFC 2867, *RADIUS Accounting Modifications for Tunnel Protocol Support*
- RFC 2868, *RADIUS Attributes for Tunnel Protocol Support*
- RFC 2869, *RADIUS Extensions*
- RFC 1492, *An Access Control Protocol, Sometimes Called TACACS*

RADIUS attributes

Commonly used standard RADIUS attributes

No.	Attribute	Description
1	User-Name	Name of the user to be authenticated.
2	User-Password	User password for PAP authentication, present only in Access-Request packets in PAP authentication mode.
3	CHAP-Password	Digest of the user password for CHAP authentication, present only in Access-Request packets in CHAP authentication mode.
4	NAS-IP-Address	IP address for the server to identify a client. Usually, a client is identified by the IP address of the access interface on the NAS, namely the NAS IP address. This attribute is present in only Access-Request packets.
5	NAS-Port	Physical port of the NAS that the user accesses.
6	Service-Type	Type of service that the user has requested or type of service to be provided.
7	Framed-Protocol	Encapsulation protocol for framed access.
8	Framed-IP-Address	IP address assigned to the user.
11	Filter-ID	Name of the filter list.
12	Framed-MTU	Maximum transmission unit (MTU) for the data link between the user and NAS. For example, with 802.1X EAP authentication, NAS uses this attribute to notify the server of the MTU for EAP packets, so as to avoid oversized EAP packets.
14	Login-IP-Host	IP address of the NAS interface that the user accesses.
15	Login-Service	Type of the service that the user uses for login.
18	Reply-Message	Text to be displayed to the user, which can be used by the server to indicate, for example, the reason of the authentication failure.
26	Vendor-Specific	Vendor specific attribute. A packet can contain one or more such proprietary attributes, each of which can contain one or more sub-attributes.
27	Session-Timeout	Maximum duration of service to be provided to the user before termination of the session.
28	Idle-Timeout	Maximum idle time permitted for the user before termination of the session.
31	Calling-Station-Id	User identification that the NAS sends to the server. For the LAN access service provided by an HP device, this attribute carries the MAC address of the user in the format HHHH-HHHH-HHHH.
32	NAS-Identifier	Identification that the NAS uses for indicating itself.

No.	Attribute	Description
40	Acct-Status-Type	Type of the Accounting-Request packet. Possible values are as follows: <ul style="list-style-type: none"> • 1—Start. • 2—Stop. • 3—Interim-Update. • 4—Reset-Charge. • 7—Accounting-On. (Defined in 3GPP, the 3rd Generation Partnership Project.) • 8—Accounting-Off. (Defined in 3GPP.) • 9 to 14—Reserved for tunnel accounting. • 15—Reserved for failed.
45	Acct-Authentic	Authentication method used by the user. Possible values are as follows: <ul style="list-style-type: none"> • 1—RADIUS. • 2—Local. • 3—Remote.
60	CHAP-Challenge	CHAP challenge generated by the NAS for MD5 calculation during CHAP authentication.
61	NAS-Port-Type	Type of the physical port of the NAS that is authenticating the user. Possible values are as follows: <ul style="list-style-type: none"> • 15—Ethernet. • 16—Any type of ADSL. • 17—Cable (with cable for cable TV). • 201—VLAN. • 202—ATM. <p>If the port is an ATM or Ethernet one and VLANs are implemented on it, the value of this attribute is 201.</p>
79	EAP-Message	Used for encapsulating EAP packets to allow the NAS to authenticate dial-in users via EAP without having to understand the EAP protocol.
80	Message-Authenticator	Used for authentication and checking of authentication packets to prevent spoofing Access-Requests. This attribute is used when RADIUS supports EAP authentication.
87	NAS-Port-Id	String for describing the port of the NAS that is authenticating the user.

HP proprietary RADIUS sub-attributes

No.	Sub-attribute	Description
1	Input-Peak-Rate	Peak rate in the direction from the user to the NAS, in bps.
2	Input-Average-Rate	Average rate in the direction from the user to the NAS, in bps.
3	Input-Basic-Rate	Basic rate in the direction from the user to the NAS, in bps.
4	Output-Peak-Rate	Peak rate in the direction from the NAS to the user, in bps.
5	Output-Average-Rate	Average rate in the direction from the NAS to the user, in bps.
6	Output-Basic-Rate	Basic rate in the direction from the NAS to the user, in bps.
15	Remanent_Volume	Remaining, available total traffic of the connection, in different units for different server types.

No.	Sub-attribute	Description
20	Command	<p>Operation for the session, used for session control. It can be:</p> <ul style="list-style-type: none"> • 1—Trigger-Request. • 2—Terminate-Request. • 3—SetPolicy. • 4—Result. • 5—PortalClear.
24	Control_Identifier	<p>Identification for retransmitted packets. For retransmitted packets of the same session, this attribute must take the same value. For retransmitted packets of different sessions, this attribute may take the same value. The client response of a retransmitted packet must also carry this attribute and the value of the attribute must be the same.</p> <p>For Accounting-Request packets of the start, stop, and interim update types, the Control-Identifier attribute, if present, makes no sense.</p>
25	Result_Code	Result of the Trigger-Request or SetPolicy operation. A value of zero means the operation succeeded. Any other value means the operation failed.
26	Connect_ID	Index of the user connection.
28	Ftp_Directory	<p>Working directory of the FTP user.</p> <p>For an FTP user, when the RADIUS client acts as the FTP server, this attribute is used to set the FTP directory on the RADIUS client.</p>
29	Exec_Privilege	Priority of the EXEC user.
59	NAS_Startup_Timestamp	Startup time of the NAS in seconds, which is represented by the time elapsed after 00:00:00 on Jan. 1, 1970 (UTC).
60	Ip_Host_Addr	User IP address and MAC address carried in authentication and accounting requests, in the format A.B.C.D hh:hh:hh:hh:hh:hh. A space is required between the IP address and the MAC address.
61	User_Notify	Information to be sent from the server to the client transparently.
62	User_HeartBeat	Hash value assigned after an 802.1X user passes authentication, which is a 32-byte string. This attribute is stored in the user list on the device and is used for verifying the handshake messages from the 802.1X user. This attribute exists in only Access-Accept and Accounting-Request packets.
140	User_Group	User groups assigned after the SSL VPN user passes authentication. A user may belong to more than one user group. In this case, the user groups are delimited by semi-colons. This attribute is used for cooperation with the SSL VPN device.
141	Security_Level	Security level assigned after the SSL VPN user passes security authentication.
201	Input-Interval-Octets	Bytes input within a real-time accounting interval.
202	Output-Interval-Octets	Bytes output within a real-time accounting interval.
203	Input-Interval-Packets	Packets input within an accounting interval, in the unit set on the device.
204	Output-Interval-Packets	Packets output within an accounting interval, in the unit set on the device.
205	Input-Interval-Gigawords	Result of bytes input within an accounting interval divided by 4G bytes.
206	Output-Interval-Gigawords	Result of bytes output within an accounting interval divided by 4G bytes.

No.	Sub-attribute	Description
207	Backup-NAS-IP	Backup source IP address for sending RADIUS packets.
255	Product_ID	Product name.

AAA configuration considerations and task list

To configure AAA, you must complete these tasks on the NAS:

1. Configure the required AAA schemes.
 - **Local authentication**—Configure local users and the related attributes, including the usernames and passwords of the users to be authenticated.
 - **Remote authentication**—Configure the required RADIUS and HWTACACS schemes. You must configure user attributes on the servers accordingly.
2. Configure AAA methods for the users' ISP domains.
 - **Authentication method**—No authentication (**none**), local authentication (**local**), or remote authentication (**scheme**)
 - **Authorization method**—No authorization (**none**), local authorization (**local**), or remote authorization (**scheme**)
 - **Accounting method**—No accounting (**none**), local accounting (**local**), or remote accounting (**scheme**)

Figure 10 AAA configuration diagram

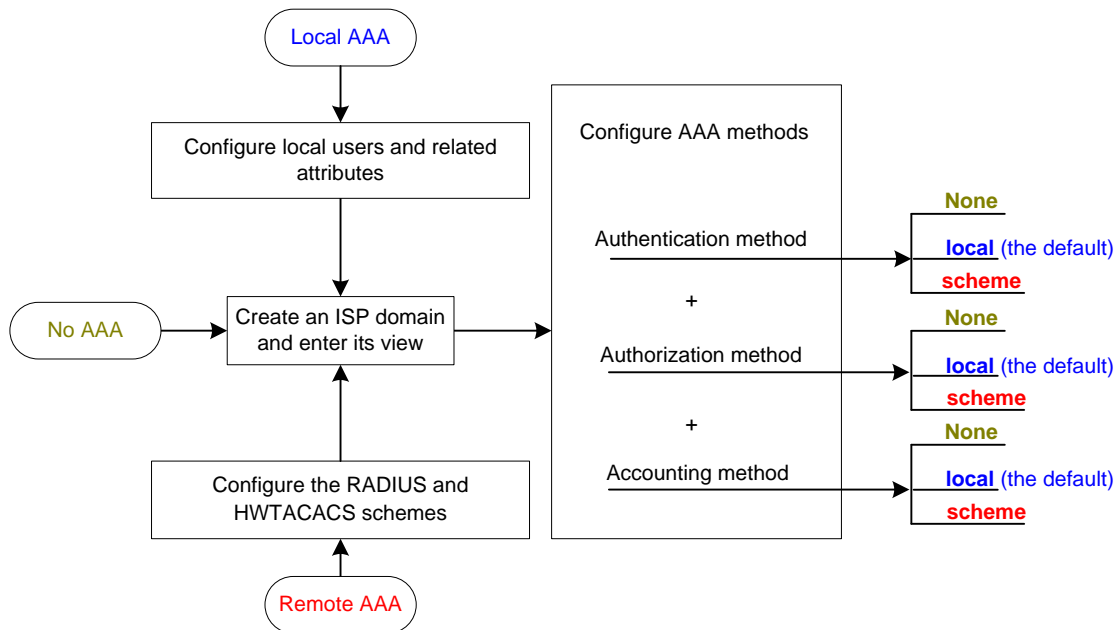


Table 4 AAA configuration task list

Task	Remarks
Configuring AAA schemes	Configuring local users Configuring RADIUS schemes
	Required. Complete at least one task.

Task	Remarks	
	Configuring HWTACACS schemes	
Configuring AAA methods for ISP domains	Creating an ISP domain	Required.
	Configuring ISP domain attributes	Optional.
	Configuring AAA authentication methods for an ISP domain	Required. Complete at least one task.
	Configuring AAA authorization methods for an ISP domain	
	Configuring AAA accounting methods for an ISP domain	
Tearing down user connections	Optional.	
Configuring a NAS ID-VLAN binding	Optional.	
Specifying the device ID used in stateful failover mode	Optional.	
Configuring a switch as a RADIUS server	Optional.	

NOTE:

To use AAA methods to control access of login users, you must configure the user interfaces to use AAA by using the **authentication-mode** command. For more information about the configuration command, see *Fundamentals Command Reference*.

Configuring AAA schemes

Configuring local users

To implement local user authentication, authorization, and accounting, you must create local users and configure user attributes on the switch. The local users and attributes are stored in the local user database on the switch. A local user is uniquely identified by a username. Configurable local user attributes are as follows:

- Service type:
Types of services that the user can use. Local authentication checks the service types of a local user. If none of the service types is available, the user cannot pass authentication.
Service types include FTP, LAN access, portal, SSH, Telnet, terminal, and Web.
- User state:
Indicates whether or not a local user can request network services. There are two user states: active and blocked. A user in active state can request network services, but a user in blocked state cannot.
- Maximum number of users using the same local user account:
Indicates how many users can use the same local user account for local authentication.
- Validity time and expiration time:
Indicates the validity time and expiration time of a local user account. A user must use a valid local user account to pass local authentication. For temporary network access requirements, you can

create a guest account and specify a validity time and an expiration time for the account to control the validity of the account.

- User group:

Each local user belongs to a local user group and bears all attributes of the group, such as the password control attributes and authorization attributes. For more information about local user group, see "[Configuring user group attributes](#)."

- Password control attributes:

Password control attributes help you control the security of local users' passwords. Password control attributes include password aging time, minimum password length, and password composition policy.

You can configure a password control attribute in system view, user group view, or local user view, making the attribute effective for all local users, all local users in a group, or only the local user. A password control attribute with a smaller effective range has a higher priority. For more information about password management and global password configuration, see "[Configuring password control](#)."

- Binding attributes:

Binding attributes are used to control the scope of users. They are checked during local authentication of a user. If the attributes of a user do not match the binding attributes configured for the local user account, the user cannot pass authentication. Binding attributes include the ISDN calling number, IP address, access port, MAC address, and native VLAN. For more information about binding attributes, see "[Configuring local user attributes](#)." Be cautious when deciding which binding attributes to configure for a local user.

- Authorization attributes:

Authorization attributes indicate the rights that a user has after passing local authentication. Authorization attributes include the ACL, PPP callback number, idle cut function, user level, user role, user profile, VLAN, and FTP/SFTP work directory. For more information about authorization attributes, see "[Configuring local user attributes](#)."

Every configurable authorization attribute has its definite application environments and purposes. When you configure authorization attributes for a local user, consider which attributes are needed and which are not.

You can configure an authorization attribute in user group view or local user view to make the attribute effective for all local users in the group or only for the local user. The setting of an authorization attribute in local user view takes precedence over that in user group view.

Local user configuration task list

Task	Remarks
Configuring local user attributes	Required
Configuring user group attributes	Optional
Displaying and maintaining local users and local user groups	Optional

Configuring local user attributes

Step	Command	Remarks
1. Enter system view.	<code>system-view</code>	N/A

Step	Command	Remarks
2. Add a local user and enter local user view.	local-user <i>user-name</i>	No local user exists by default.
3. Configure a password for the local user.	password [{ cipher simple } <i>password</i>]	Optional. A local user with no password configured directly passes authentication after providing the valid local username and attributes. To enhance security, configure a password for each local user. If none of the parameters is specified, you enter the interactive mode to set a plaintext password. This interactive mode is supported only on switches that support the password control feature.
4. Specify the service types for the local user.	service-type { ftp lan-access { ssh telnet terminal } * portal web }	By default, no service is authorized to a local user.
5. Place the local user to the state of active or blocked.	state { active block }	Optional. When created, a local user is in active state by default, and the user can request network services.
6. Set the maximum number of concurrent users of the local user account.	access-limit <i>max-user-number</i>	Optional. By default, there is no limit to the maximum number of concurrent users of a local user account. The limit is effective only for local accounting, and is not effective for FTP users.
7. Configure the password control attributes for the local user.	<ul style="list-style-type: none"> Set the password aging time: password-control aging <i>aging-time</i> Set the minimum password length: password-control length <i>length</i> Configure the password composition policy: password-control composition <i>type-number type-number</i> [<i>type-length type-length</i>] 	Optional. By default, the password control attributes of the user group to which the local user belongs apply, and any password control attribute that is not configured in the user group uses the global setting. The global settings include a 90-day password aging time, a minimum password length of 10 characters, and at least one password composition type and at least one character required for each password composition type.
8. Configure the binding attributes for the local user.	bind-attribute { call-number <i>call-number</i> [: <i>subcall-number</i>] ip <i>ip-address</i> location port <i>slot-number subslot-number port-number</i> mac <i>mac-address</i> vlan <i>vlan-id</i> } *	Optional. By default, no binding attribute is configured for a local user. Binding attributes are only intended for and LAN users.

Step	Command	Remarks
9. Configure the authorization attributes for the local user.	authorization-attribute { acl <i>acl-number</i> callback-number <i>callback-number</i> idle-cut <i>minute</i> level <i>level</i> user-profile <i>profile-name</i> user-role { guest guest-manager security-audit } vlan <i>vlan-id</i> work-directory <i>directory-name</i> } *	Optional. By default, no authorization attribute is configured for a local user. For LAN and portal users, only acl , idle-cut , user-profile , and vlan are supported. For SSH, terminal, and Web users, only level is supported. For FTP users, only level and work-directory are supported. For Telnet users, only level and user-role is supported. For other types of local users, no binding attribute is supported.
10. Set the validity time of the local user.	validity-date <i>time</i>	Optional. Not set by default.
11. Set the expiration time of the local user.	expiration-date <i>time</i>	Optional. Not set by default.
12. Assign the local user to a user group.	group <i>group-name</i>	Optional. By default, a local user belongs to the default user group system .

- For more information about password control configuration commands, see *Security Command Reference*.
- If the user interface authentication mode (set by the **authentication-mode** command in user interface view) is AAA (**scheme**), which commands a login user can use after login depends on the privilege level authorized to the user. If the user interface authentication mode is password (**password**) or no authentication (**none**), which commands a login user can use after login depends on the level configured for the user interface (set by the **user privilege level** command in user interface view). For an SSH user using public key authentication, which commands are available depends on the level configured for the user interface. For more information about user interface authentication mode and user interface command level, see *Fundamentals Configuration Guide*.
- You can configure the user profile authorization attribute in local user view, user group view, and ISP domain view. The setting in local user view has the highest priority, and that in ISP domain view has the lowest priority. For more information about user profiles, see "[Configuring a user profile](#)."
- You cannot delete a local user who is the only security log manager in the system, nor can you change or delete the security log manager role of the user. To do so, you must specify a new security log manager first.

Configuring user group attributes

User groups simplify local user configuration and management. A user group consists of a group of local users and has a set of local user attributes. You can configure local user attributes for a user group to implement centralized user attributes management for the local users in the group. Configurable user attributes include password control attributes and authorization attributes.

By default, every newly added local user belongs to the system default user group system and bears all attributes of the group. To change the user group to which a local user belongs, use the **user-group** command in local user view.

To configure attributes for a user group:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a user group and enter user group view.	user-group <i>group-name</i>	N/A
3. Configure password control attributes for the user group.	<ul style="list-style-type: none"> Set the password aging time: password-control aging <i>aging-time</i> Set the minimum password length: password-control length <i>length</i> Configure the password composition policy: password-control composition <i>type-number type-number [type-length type-length]</i> 	Optional. By default, the global settings apply. The global settings include a 90-day password aging time, a minimum password length of 10 characters, and at least one password composition type and at least one character required for each password composition type.
4. Configure the authorization attributes for the user group.	authorization-attribute { acl <i>acl-number</i> callback-number <i>callback-number</i> idle-cut <i>minute</i> level <i>level</i> user-profile <i>profile-name</i> vlan <i>vlan-id</i> work-directory <i>directory-name</i> } *	Optional. By default, no authorization attribute is configured for a user group.
5. Set the guest attribute for the user group.	group-attribute allow-guest	Optional. By default, the guest attribute is not set for a user group, and guest users created by a guest manager through the Web interface cannot join the group.

NOTE:

For more information about password control attributes configuration commands, see *Security Command Reference*.

Displaying and maintaining local users and local user groups

Task	Command	Remarks
Display local user information	display local-user [idle-cut { disable enable } service-type { ftp lan-access portal ssh telnet terminal web } state { active block } user-name <i>user-name</i> vlan <i>vlan-id</i>] [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display the user group configuration information.	display user-group [<i>group-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

Configuring RADIUS schemes

A RADIUS scheme specifies the RADIUS servers that the switch can cooperate with and defines a set of parameters that the switch uses to exchange information with the RADIUS servers. There may be authentication/authorization servers and accounting servers, or primary servers and secondary servers. The parameters include the IP addresses of the servers, the shared keys, and the RADIUS server type.

RADIUS scheme configuration task list

Task	Remarks
Creating a RADIUS scheme	Required
Specifying the RADIUS authentication/authorization servers	Required
Specifying the RADIUS accounting servers and the relevant parameters	Optional
Specifying the shared keys for secure RADIUS communication	Optional
Specifying the VPN to which the servers belong	Optional
Setting the username format and traffic statistics units	Optional
Setting the supported RADIUS server type	Optional
Setting the maximum number of RADIUS request transmission attempts	Optional
Setting the status of RADIUS servers	Optional
Specifying the source IP address for outgoing RADIUS packets	Optional
Specifying a backup source IP address for outgoing RADIUS packets	Optional
Setting timers for controlling communication with RADIUS servers	Optional
Configuring RADIUS accounting-on	Optional
Configuring the IP address of the security policy server	Optional
Configuring interpretation of RADIUS class attribute as CAR parameters	Optional
Enabling the trap function for RADIUS	Optional
Enabling the RADIUS listening port of the RADIUS client	Optional
Setting the DSCP value for RADIUS protocol packets	Optional
Displaying and maintaining RADIUS	Optional

Creating a RADIUS scheme

Before performing other RADIUS configurations, follow these steps to create a RADIUS scheme and enter RADIUS scheme view:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a RADIUS scheme and enter RADIUS scheme view.	radius scheme <i>radius-scheme-name</i>	No RADIUS scheme exists by default.

NOTE:

A RADIUS scheme can be referenced by multiple ISP domains at the same time.

Specifying the RADIUS authentication/authorization servers

You can specify one primary authentication/authorization server and up to 16 secondary authentication/authorization servers for a RADIUS scheme. When the primary server is not available, a secondary server is used. In a scenario where redundancy is not required, specify only the primary server.

In RADIUS, user authorization information is piggybacked in authentication responses sent to RADIUS clients. There is no separate RADIUS authorization server.

You can enable the server status detection feature. With the feature, the switch periodically sends an authentication request to check whether or not the target RADIUS authentication/authorization server is reachable. If yes, the switch sets the status of the server to **active**. If not, the switch sets the status of the server to **block**. This feature can promptly notify authentication modules of latest server status information. For example, server status detection can work with the 802.1X critical VLAN feature, so that the switch can trigger 802.1X authentication for users in the critical VLAN immediately on detection of a reachable RADIUS authentication/authorization server.

Follow these guidelines when you specify RADIUS authentication/authorization servers:

- The IP addresses of the primary and secondary authentication/authorization servers for a scheme must be different from each other. Otherwise, the configuration fails.
- All servers for authentication/authorization and accounting, primary or secondary, must use IP addresses of the same IP version.
- You can specify a RADIUS authentication/authorization server as the primary authentication/authorization server for one scheme and as the secondary authentication/authorization server for another scheme at the same time.

To specify RADIUS authentication/authorization servers for a RADIUS scheme:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RADIUS scheme view.	radius scheme <i>radius-scheme-name</i>	N/A

Step	Command	Remarks
3. Specify RADIUS authentication/authorization servers.	<ul style="list-style-type: none"> Specify the primary RADIUS authentication/authorization server: primary authentication { <i>ip-address</i> ipv6 <i>ipv6-address</i> } [<i>port-number</i> key [cipher simple] <i>key</i> probe username <i>name</i> [interval <i>interval</i>] vpn-instance <i>vpn-instance-name</i>] * Specify a secondary RADIUS authentication/authorization server: secondary authentication { <i>ip-address</i> ipv6 <i>ipv6-address</i> } [<i>port-number</i> key [cipher simple] <i>key</i> probe username <i>name</i> [interval <i>interval</i>] vpn-instance <i>vpn-instance-name</i>] * 	<p>Configure at least one command.</p> <p>No authentication/authorization server is specified by default.</p>

Specifying the RADIUS accounting servers and the relevant parameters

You can specify one primary accounting server and up to 16 secondary accounting servers for a RADIUS scheme. When the primary server is not available, a secondary server is used. When redundancy is not required, specify only the primary server.

By setting the maximum number of real-time accounting attempts for a scheme, you make the switch disconnect users for whom no accounting response is received before the number of accounting attempts reaches the limit.

When the switch receives a connection teardown request from a host or a connection teardown notification from an administrator, it sends a stop-accounting request to the accounting server. You can enable buffering of non-responded stop-accounting requests to allow the switch to buffer and resend a stop-accounting request until it receives a response or the number of stop-accounting attempts reaches the configured limit. In the latter case, the switch discards the packet.

Follow these guidelines when you specify RADIUS accounting servers:

- The IP addresses of the primary and secondary accounting servers must be different from each other. Otherwise, the configuration fails.
- All servers for authentication/authorization and accountings, primary or secondary, must use IP addresses of the same IP version.
- If you delete an accounting server that is serving users, the switch can no longer send real-time accounting requests and stop-accounting requests for the users to that server, or buffer the stop-accounting requests.
- You can specify a RADIUS accounting server as the primary accounting server for one scheme and as the secondary accounting server for another scheme at the same time.
- RADIUS does not support accounting for FTP users.

To specify RADIUS accounting servers and set relevant parameters for a scheme:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RADIUS scheme view.	radius scheme <i>radius-scheme-name</i>	N/A

Step	Command	Remarks
3. Specify RADIUS accounting servers.	<ul style="list-style-type: none"> Specify the primary RADIUS accounting server: primary accounting { <i>ip-address</i> ipv6 <i>ipv6-address</i> } [<i>port-number</i> key [cipher simple] <i>key</i> vpn-instance <i>vpn-instance-name</i>] * Specify a secondary RADIUS accounting server: secondary accounting { <i>ip-address</i> ipv6 <i>ipv6-address</i> } [<i>port-number</i> key [cipher simple] <i>key</i> vpn-instance <i>vpn-instance-name</i>] * 	<p>Configure at least one command.</p> <p>No accounting server is specified by default.</p>
4. Set the maximum number of real-time accounting attempts.	retry realtime-accounting <i>retry-times</i>	<p>Optional.</p> <p>The default setting is 5.</p>
5. Enable buffering of stop-accounting requests to which no responses are received.	stop-accounting-buffer enable	<p>Optional.</p> <p>Enabled by default.</p>
6. Set the maximum number of stop-accounting attempts.	retry stop-accounting <i>retry-times</i>	<p>Optional.</p> <p>The default setting is 500.</p>

Specifying the shared keys for secure RADIUS communication

The RADIUS client and RADIUS server use the MD5 algorithm to authenticate packets exchanged between them and use shared keys for packet authentication and user passwords encryption. They must use the same key for the same type of communication.

A shared key configured in this task is for all servers of the same type (accounting or authentication) in the scheme, and has a lower priority than a shared key configured individually for a RADIUS server.

To specify a shared key for secure RADIUS communication:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RADIUS scheme view.	radius scheme <i>radius-scheme-name</i>	N/A
3. Specify a shared key for secure RADIUS authentication/authorization or accounting communication.	key { accounting authentication } <i>key</i>	No shared key is specified by default.

NOTE:

A shared key configured on the switch must be the same as that configured on the RADIUS server.

Specifying the VPN to which the servers belong (available only on the HP 5500 EI)

After you specify a VPN for a RADIUS scheme, all the authentication/authorization/accounting servers specified for the scheme belong to the VPN. However, if you also specify a VPN when specifying a server for the scheme, the server belongs to the specific VPN.

To specify a VPN for a RADIUS scheme:

Step	Command
1. Enter system view.	system-view
2. Enter RADIUS scheme view.	radius scheme <i>radius-scheme-name</i>
3. Specify a VPN for the RADIUS scheme.	vpn-instance <i>vpn-instance-name</i>

Setting the username format and traffic statistics units

A username is usually in the format of *userid@isp-name*, where *isp-name* represents the name of the ISP domain the user belongs to and is used by the switch to determine which users belong to which ISP domains. However, some earlier RADIUS servers cannot recognize usernames that contain an ISP domain name. In this case, the switch must remove the domain name of each username before sending the username. You can set the username format on the switch for this purpose.

The switch periodically sends accounting updates to RADIUS accounting servers to report the traffic statistics of online users. For normal and accurate traffic statistics, make sure the unit for data flows and that for packets on the switch are consistent with those on the RADIUS server.

Follow these guidelines when you set the username format and the traffic statistics units for a RADIUS scheme:

- If a RADIUS scheme defines that the username is sent without the ISP domain name, do not apply the RADIUS scheme to more than one ISP domain. Otherwise, users using the same username but in different ISP domains are considered the same user.
- For level switching authentication, the **user-name-format keep-original** and **user-name-format without-domain** commands produce the same results. They make sure usernames sent to the RADIUS server carry no ISP domain name.

To set the username format and the traffic statistics units for a RADIUS scheme:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RADIUS scheme view.	radius scheme <i>radius-scheme-name</i>	N/A
3. Set the format for usernames sent to the RADIUS servers.	user-name-format { keep-original with-domain without-domain }	Optional. By default, the ISP domain name is included in a username.
4. Specify the unit for data flows or packets sent to the RADIUS servers.	data-flow-format { data { byte giga-byte kilo-byte mega-byte } packet { giga-packet kilo-packet mega-packet one-packet } }*	Optional. The default unit is byte for data flows and is one-packet for data packets.

Setting the supported RADIUS server type

The supported RADIUS server type determines the type of the RADIUS protocol that the switch uses to communicate with the RADIUS server. It can be standard or extended:

- **Standard**—Uses the standard RADIUS protocol, compliant to RFC 2865 and RFC 2866 or later.
- **Extended**—Uses the proprietary RADIUS protocol of HP.

When the RADIUS server runs on IMC, you must set the RADIUS server type to **extended**. When the RADIUS server runs third-party RADIUS server software, either RADIUS server type applies. For the switch to function as a RADIUS server to authenticate login users, you must set the RADIUS server type to **standard**.

To set the RADIUS server type:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RADIUS scheme view.	radius scheme <i>radius-scheme-name</i>	N/A
3. Set the RADIUS server type.	server-type { extended standard }	Optional. The default RADIUS server type is standard .

NOTE:

Changing the RADIUS server type restores the unit for data flows and that for packets that are sent to the RADIUS server to the defaults.

Setting the maximum number of RADIUS request transmission attempts

Because RADIUS uses UDP packets to transfer data, the communication process is not reliable. RADIUS uses a retransmission mechanism to improve the reliability. If a NAS sends a RADIUS request to a RADIUS server but receives no response after the response timeout timer (defined by the **timer response-timeout** command) expires, it retransmits the request. If the number of transmission attempts exceeds the specified limit but it still receives no response, it tries to communicate with other RADIUS servers in active state. If no other servers are in active state at the time, it considers the authentication or accounting attempt a failure. For more information about RADIUS server states, see "[Setting the status of RADIUS servers.](#)"

To set the maximum number of RADIUS request transmission attempts for a scheme:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RADIUS scheme view.	radius scheme <i>radius-scheme-name</i>	N/A
3. Set the maximum number of RADIUS request transmission attempts.	retry <i>retry-times</i>	Optional. The default setting is 3.

NOTE:

- The maximum number of transmission attempts of RADIUS packets multiplied by the RADIUS server response timeout period cannot be greater than 75 seconds.
- For more information about the RADIUS server response timeout period, see "[Setting timers for controlling communication with RADIUS servers.](#)"

Setting the status of RADIUS servers

By setting the status of RADIUS servers to blocked or active, you can control which servers the switch communicates with for authentication, authorization, and accounting or turn to when the current servers

are not available anymore. In practice, you can specify one primary RADIUS server and multiple secondary RADIUS servers, with the secondary servers functioning as the backup of the primary servers. Generally, the switch chooses servers based on these rules:

- When the primary server is in active state, the switch communicates with the primary server. If the primary server fails, the switch changes the server's status to blocked and starts a quiet timer for the server, and then turns to a secondary server in active state (a secondary server configured earlier has a higher priority). If the secondary server is unreachable, the switch changes the server's status to blocked, starts a quiet timer for the server, and continues to check the next secondary server in active state. This search process continues until the switch finds an available secondary server or has checked all secondary servers in active state. If the quiet timer of a server expires or an authentication or accounting response is received from the server, the status of the server changes back to active automatically, but the switch does not check the server again during the authentication or accounting process. If no server is found reachable during one search process, the switch considers the authentication or accounting attempt a failure.
- Once the accounting process of a user starts, the switch keeps sending the user's real-time accounting requests and stop-accounting requests to the same accounting server. If you remove the accounting server, real-time accounting requests and stop-accounting requests for the user cannot be delivered to the server anymore.
- If you remove an authentication or accounting server in use, the communication of the switch with the server soon times out, and the switch looks for a server in active state from scratch by checking any primary server first and then secondary servers in the order they are configured.
- When the primary server and secondary servers are all in blocked state, the switch communicates with the primary server. If the primary server is available, its status changes to active. Otherwise, its status remains to be blocked.
- If one server is in active state and all the others are in blocked state, the switch only tries to communicate with the server in active state, even if the server is unavailable.
- After receiving an authentication/accounting response from a server, the switch changes the status of the server identified by the source IP address of the response to active if the current status of the server is blocked.

By default, the switch sets the status of all RADIUS servers to active. In cases such as a server failure, you can change the status of the server to blocked to avoid communication with the server.

To set the status of RADIUS servers in a RADIUS scheme:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RADIUS scheme view.	radius scheme <i>radius-scheme-name</i>	N/A

Step	Command	Remarks
3. Set the RADIUS server status.	<ul style="list-style-type: none"> Set the status of the primary RADIUS authentication/authorization server: state primary authentication { active block } Set the status of the primary RADIUS accounting server: state primary accounting { active block } Set the status of a secondary RADIUS authentication/authorization server: state secondary authentication [ip ipv4-address ipv6 ipv6-address] { active block } Set the status of a secondary RADIUS accounting server: state secondary accounting [ip ipv4-address ipv6 ipv6-address] { active block } 	Optional. By default, all servers in the RADIUS scheme are in active state.

NOTE:

- The server status set by the **state** command cannot be saved to the configuration file. After the switch restarts, the status of each server is restored to active.
- To display the states of the servers, use the **display radius scheme** command.

Specifying the source IP address for outgoing RADIUS packets

The source IP address of RADIUS packets that a NAS sends must match the IP address of the NAS configured on the RADIUS server. A RADIUS server identifies a NAS by its IP address. Upon receiving a RADIUS packet, a RADIUS server checks whether the source IP address of the packet is the IP address of any managed NAS. If yes, the server processes the packet. If not, the server drops the packet.

Usually, the source address of outgoing RADIUS packets can be the IP address of the NAS's any interface that can communicate with the RADIUS server. In some special scenarios, however, you must change the source IP address. For example, if a Network Address Translation (NAT) device is present between the NAS and the RADIUS server, the source IP address of outgoing RADIUS packets must be a public IP address of the NAS. If the NAS is configured with the Virtual Router Redundancy Protocol (VRRP) for stateful failover, the source IP address of outgoing RADIUS packets can be the virtual IP address of the VRRP group to which the uplink belongs.

You can specify a source IP address for outgoing RADIUS packets in RADIUS scheme view for a specific RADIUS scheme, or in system view for all RADIUS schemes whose servers are in a VPN or the public network. Before sending a RADIUS packet, a NAS selects a source IP address in this order:

- The source IP address specified for the RADIUS scheme.
- The source IP address specified in system view for the VPN or public network, depending on where the RADIUS server resides.
- The IP address of the outbound interface specified by the route.

To specify a source IP address for all RADIUS schemes in a VPN or the public network:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Specify a source IP address for outgoing RADIUS packets.	radius nas-ip { <i>ip-address</i> ipv6 <i>ipv6-address</i> } [vpn-instance <i>vpn-instance-name</i>]	By default, the IP address of the outbound interface is used as the source IP address.

To specify a source IP address for a specific RADIUS scheme:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RADIUS scheme view.	radius scheme <i>radius-scheme-name</i>	N/A
3. Specify a source IP address for outgoing RADIUS packets.	nas-ip { <i>ip-address</i> ipv6 <i>ipv6-address</i> }	By default, the IP address of the outbound interface is used as the source IP address.

Specifying a backup source IP address for outgoing RADIUS packets (available only on the HP 5500 EI)

In a stateful failover scenario, the active switch authenticates portal users by interacting with the RADIUS server, and synchronizes its online portal user information to the standby switch through the backup link established between them. The standby switch only receives and processes synchronization messages from the active switch. However, when the active switch fails, the RADIUS server does not send RADIUS packets to the standby switch because it does not know the IP address of the standby switch. To solve this problem, configure the source IP address for outgoing RADIUS packets on each switch as the backup source IP address for outgoing RADIUS packets on the other switch. With such configuration, the active switch sends the source IP address for outgoing RADIUS packets that is configured on the standby switch to the RADIUS server, so that the RADIUS server can send unsolicited RADIUS packets to the standby switch.

You can specify a backup IP address for outgoing RADIUS packets in RADIUS scheme view for a specific RADIUS scheme, or in system view for all RADIUS schemes whose servers are in a VPN or the public network. Before sending a RADIUS packet, a NAS selects a backup source IP address in this order:

1. The backup source IP address specified for the RADIUS scheme.
2. The backup source IP address specified in system view for the VPN or public network, depending on where the RADIUS server resides.

If no backup source IP address is specified in the views, the NAS sends no backup source IP address to the server.

To specify a backup source IP address for all RADIUS schemes of a VPN or the public network:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Specify a backup source IP address for outgoing RADIUS packets.	radius nas-backup-ip <i>ip-address</i> [vpn-instance <i>vpn-instance-name</i>]	Not specified by default.

To specify a backup source IP address for a RADIUS scheme:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RADIUS scheme view.	radius scheme <i>radius-scheme-name</i>	N/A
3. Specify a backup source IP address for outgoing RADIUS packets.	nas-backup-ip <i>ip-address</i>	Not specified by default.

NOTE:

The backup source IP address specified for outgoing RADIUS packets takes effect only when stateful failover is configured, and it must be the source IP address for outgoing RADIUS packets that is configured on the standby switch.

Setting timers for controlling communication with RADIUS servers

The switch uses the following types of timers to control the communication with a RADIUS server:

- **Server response timeout timer (response-timeout)**—Defines the RADIUS request retransmission interval. After sending a RADIUS request (authentication/authorization or accounting request), the switch starts this timer. If the switch receives no response from the RADIUS server before this timer expires, it resends the request.
- **Server quiet timer (quiet)**—Defines the duration to keep an unreachable server in blocked state. If a server is not reachable, the switch changes the server’s status to blocked, starts this timer for the server, and tries to communicate with another server in active state. After this timer expires, the switch changes the status of the server back to active.
- **Real-time accounting timer (realtime-accounting)**—Defines the interval at which the switch sends real-time accounting packets to the RADIUS accounting server for online users. To implement real-time accounting, the switch must periodically send real-time accounting packets to the accounting server for online users.

To set timers for controlling communication with RADIUS servers:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RADIUS scheme view.	radius scheme <i>radius-scheme-name</i>	N/A
3. Set the RADIUS server response timeout timer.	timer response-timeout <i>seconds</i>	Optional. The default RADIUS server response timeout timer is 3 seconds.
4. Set the quiet timer for the servers.	timer quiet <i>minutes</i>	Optional. The quiet timer is 5 minutes.
5. Set the real-time accounting timer.	timer realtime-accounting <i>minutes</i>	Optional. The default real-time accounting timer is 12 minutes.

- For a type of users, the maximum number of transmission attempts multiplied by the RADIUS server response timeout period must be less than the client connection timeout time and must not exceed

75 seconds. Otherwise, stop-accounting messages cannot be buffered, and the primary/secondary server switchover cannot take place. For example, the product of the two parameters must be less than 10 seconds for voice users, and less than 30 seconds for Telnet users because the client connection timeout period for voice users is 10 seconds and that for Telnet users is 30 seconds.

- When you configure the maximum number of RADIUS packet transmission attempts and the RADIUS server response timeout period, be sure to take the number of secondary servers into account. If the retransmission process takes too much time, the client connection in the access module may be timed out while the switch is trying to find an available server.
- When a number of secondary servers are configured, the client connections of access modules that have a short client connection timeout period may still be timed out during initial authentication or accounting, even if the packet transmission attempt limit and server response timeout period are configured with small values. In this case, the next authentication or accounting attempt may succeed because the switch has set the state of the unreachable servers to blocked and the time for finding a reachable server is shortened.
- Be sure to set the server quiet timer properly. Too short a quiet timer may result in frequent authentication or accounting failures because the switch has to repeatedly attempt to communicate with an unreachable server that is in active state.
- For more information about the maximum number of RADIUS packet transmission attempts, see "[Setting the maximum number of RADIUS request transmission attempts.](#)"

Configuring RADIUS accounting-on

The accounting-on feature enables a switch to send accounting-on packets to the RADIUS server after it reboots, making the server log out users who logged in through the switch before the reboot. Without this feature, users who were online before the reboot cannot re-log in after the reboot, because the RADIUS server considers they are already online.

If a switch sends an accounting-on packet to the RADIUS server but receives no response, it resends the packet to the server at a particular interval for a specified number of times.

To configure the accounting-on feature for a RADIUS scheme:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RADIUS scheme view.	radius scheme <i>radius-scheme-name</i>	N/A
3. Enable accounting-on and configure parameters.	accounting-on enable [interval seconds send send-times] *	Disabled by default. The default interval is 3 seconds and the default number of send-times is 50.

NOTE:

The accounting-on feature requires the cooperation of the HP IMC network management system.

Configuring the IP address of the security policy server

The core of the HP EAD solution is integration and cooperation, and the security policy server is the management and control center. Using a collection of software, the security policy server provides functions such as user management, security policy management, security status assessment, security cooperation control, and security event audit.

The NAS checks the validity of received control packets and accepts only control packets from known servers. To use a security policy server that is independent of the AAA servers, you must configure the IP address of the security policy server on the NAS. To implement all EAD functions, configure both the IP address of the IMC security policy server and that of the IMC Platform on the NAS.

To configure the IP address of the security policy server for a scheme:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RADIUS scheme view.	radius scheme <i>radius-scheme-name</i>	N/A
3. Specify a security policy server.	security-policy-server <i>ip-address</i>	No security policy server is specified by default.

Configuring interpretation of RADIUS class attribute as CAR parameters

According to RFC 2865, a RADIUS server assigns the RADIUS class attribute (attribute 25) to a RADIUS client. However, the RFC only requires the RADIUS client to send the attribute to the accounting server on an "as is" basis. It does not require the RADIUS client to interpret the attribute. Some RADIUS servers use the class attribute to deliver the assigned committed access rate (CAR) parameters. In this case, the switch must interpret the attribute as the CAR parameters to implement user-based traffic monitoring and controlling.

To configure the switch to interpret the RADIUS class attribute as CAR parameters:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter RADIUS scheme view.	radius scheme <i>radius-scheme-name</i>	N/A
3. Interpret the class attribute as CAR parameters.	attribute 25 car	By default, RADIUS attribute 25 is not interpreted as CAR parameters.

NOTE:

Whether interpretation of RADIUS class attribute as CAR parameters is supported depends on two factors:

- Whether the switch supports CAR parameters assignment.
- Whether the RADIUS server supports assigning CAR parameters through the class attribute.

Enabling the trap function for RADIUS

With the trap function, a NAS sends a trap message when either of the following events occurs:

- The status of a RADIUS server changes. If a NAS receives no response to an accounting or authentication request before the specified maximum number of RADIUS request transmission attempts is exceeded, it considers the server unreachable, sets the status of the server to **block** and sends a trap message. If the NAS receives a response from a RADIUS server that it considers unreachable, the NAS considers that the RADIUS server is reachable again, sets the status of the server to **active**, and sends a trap message.
- The ratio of the number of failed transmission attempts to the total number of authentication request transmission attempts reaches the threshold. This threshold ranges from 1% to 100% and defaults to 30%. This threshold can only be configured through the MIB.

The failure ratio is generally small. If a trap message is triggered because the failure ratio is higher than the threshold, troubleshoot the configuration on and the communication between the NAS and the RADIUS server.

To enable the trap function for RADIUS:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable the trap function for RADIUS.	radius trap { accounting-server-down authentication-error-threshold authentication-server-down }	Disabled by default.

Enabling the RADIUS listening port of the RADIUS client

Only after you enable the RADIUS listening port of a RADIUS client, can the client receive and send RADIUS packets. If RADIUS is not required, disable the RADIUS listening port to avoid attacks that exploit RADIUS packets.

To enable the RADIUS listening port of a RADIUS client:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable the RADIUS listening port of a RADIUS client.	radius client enable	Optional. Enabled by default.

Setting the DSCP value for RADIUS protocol packets

A field in an IPv4 or IPv6 header contains eight bits and is used to identify the service type of an IP packet. In an IPv4 packet, this field is called "Type of Service (ToS)." In an IPv6 packet, this field is called "Traffic class." According to RFC 2474, the ToS field is redefined as the differentiated services (DS) field, where a DSCP value is represented by the first six bits (0 to 5) and is in the range 0 to 63. The remaining two bits (6 and 7) are reserved. When a packet is being transmitted, the network devices can identify its DSCP value, and determines the transmission priority of the packet according to the DSCP value.

When you configure the DSCP value for some types of protocol packets, you should specify the ToS field value rather than the DSCP value. Because the DSCP field is the first six bits of the ToS field, each four continuous ToS field values, starting from 0, correspond to one DSCP value. An easier way to convert the DSCP value to the ToS value is to multiply the expected DSCP value by four to get the ToS field value.

To set the DSCP value for RADIUS protocol packets:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the DSCP value for IPv4 RADIUS protocol packets.	radius dscp dscp-value	Optional. By default, the DSCP value in IPv4 RADIUS protocol packets is 0.
3. Set the DSCP value for IPv6 RADIUS protocol packets.	radius ipv6 dscp dscp-value	Optional. By default, the DSCP value in IPv6 RADIUS protocol packets is 0.

Displaying and maintaining RADIUS

Task	Command	Remarks
Display the configuration information of RADIUS schemes.	display radius scheme [<i>radius-scheme-name</i>] [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the statistics for RADIUS packets .	display radius statistics [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about buffered stop-accounting requests for which no responses have been received .	display stop-accounting-buffer { radius-scheme <i>radius-server-name</i> session-id <i>session-id</i> time-range <i>start-time stop-time</i> user-name <i>user-name</i> } [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear RADIUS statistics .	reset radius statistics [slot <i>slot-number</i>]	Available in user view
Clear the buffered stop-accounting requests for which no responses have been receive.	reset stop-accounting-buffer { radius-scheme <i>radius-server-name</i> session-id <i>session-id</i> time-range <i>start-time stop-time</i> user-name <i>user-name</i> } [slot <i>slot-number</i>]	Available in user view

Configuring HWTACACS schemes

NOTE:

You cannot remove the HWTACACS schemes in use or change the IP addresses of the HWTACACS servers in use.

HWTACACS configuration task list

Task	Remarks
Creating an HWTACACS scheme	Required
Specifying the HWTACACS authentication servers	Required
Specifying the HWTACACS authorization servers	Optional
Specifying the HWTACACS accounting servers and the relevant parameters	Optional
Specifying the shared keys for secure HWTACACS communication	Required
Specifying the VPN to which the servers belong	Optional
Setting the username format and traffic statistics units	Optional
Specifying a source IP address for outgoing HWTACACS packets	Optional
Setting timers for controlling communication with HWTACACS servers	Optional
Displaying and maintaining HWTACACS	Optional

Creating an HWTACACS scheme

The HWTACACS protocol is configured on a per scheme basis. Before performing other HWTACACS configurations, follow these steps to create an HWTACACS scheme and enter HWTACACS scheme view:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create an HWTACACS scheme and enter HWTACACS scheme view.	hwtacacs scheme <i>hwtacacs-scheme-name</i>	Not defined by default.

NOTE:

- Up to 16 HWTACACS schemes can be configured.
- A scheme can be deleted only when it is not referenced.

Specifying the HWTACACS authentication servers

You can specify one primary authentication server and up to one secondary authentication server for an HWTACACS scheme. When the primary server is not available, any secondary server is used. In a scenario where redundancy is not required, specify only the primary server.

Follow these guidelines when you specify HWTACACS authentication servers:

- An HWTACACS server can function as the primary authentication server of one scheme and as the secondary authentication server of another scheme at the same time.
- The IP addresses of the primary and secondary authentication servers cannot be the same. Otherwise, the configuration fails.
- You can remove an authentication server only when no active TCP connection for sending authentication packets is using it.

To specify HWTACACS authentication servers for an HWTACACS scheme:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter HWTACACS scheme view.	hwtacacs scheme <i>hwtacacs-scheme-name</i>	N/A
3. Specify HWTACACS authentication servers.	<ul style="list-style-type: none">• Specify the primary HWTACACS authentication server: primary authentication <i>ip-address</i> [<i>port-number</i> vpn-instance <i>vpn-instance-name</i>] *• Specify the secondary HWTACACS authentication server: secondary authentication <i>ip-address</i> [<i>port-number</i> vpn-instance <i>vpn-instance-name</i>] *	<ul style="list-style-type: none">• Configure at least one command.• No authentication server is specified by default.

Specifying the HWTACACS authorization servers

You can specify one primary authorization server and up to one secondary authorization server for an HWTACACS scheme. When the primary server is not available, any secondary server is used. In a scenario where redundancy is not required, specify only the primary server.

Follow these guidelines when you specify HWTACACS authorization servers:

- An HWTACACS server can function as the primary authorization server of one scheme and as the secondary authorization server of another scheme at the same time.
- The IP addresses of the primary and secondary authorization servers cannot be the same. Otherwise, the configuration fails.
- You can remove an authorization server only when no active TCP connection for sending authorization packets is using it.

To specify HWTACACS authorization servers for an HWTACACS scheme:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter HWTACACS scheme view.	hwtacacs scheme <i>hwtacacs-scheme-name</i>	N/A
3. Specify HWTACACS authorization servers.	<ul style="list-style-type: none">• Specify the primary HWTACACS authorization server: primary authorization <i>ip-address</i> [<i>port-number</i> vpn-instance <i>vpn-instance-name</i>] *• Specify the secondary HWTACACS authorization server: secondary authorization <i>ip-address</i> [<i>port-number</i> vpn-instance <i>vpn-instance-name</i>] *	Configure at least one command. No authorization server is specified by default.

Specifying the HWTACACS accounting servers and the relevant parameters

You can specify one primary accounting server and up to one secondary accounting server for an HWTACACS scheme. When the primary server is not available, any secondary server is used. In a scenario where redundancy is not required, specify only the primary server.

When the switch receives a connection teardown request from a host or a connection teardown command from an administrator, it sends a stop-accounting request to the accounting server. You can enable buffering of non-responded stop-accounting requests to allow the switch to buffer and resend a stop-accounting request until it receives a response or the number of stop-accounting attempts reaches the configured limit. In the latter case, the switch discards the packet.

Follow these guidelines when you specify HWTACACS accounting servers:

- An HWTACACS server can function as the primary accounting server of one scheme and as the secondary accounting server of another scheme at the same time.
- The IP addresses of the primary and secondary accounting servers cannot be the same. Otherwise, the configuration fails.
- You can remove an accounting server only when no active TCP connection for sending accounting packets is using it.
- HWTACACS does not support accounting for FTP users.

To specify HWTACACS accounting servers and set relevant parameters for an HWTACACS scheme:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter HWTACACS scheme view.	hwtacacs scheme <i>hwtacacs-scheme-name</i>	N/A
3. Specify HWTACACS accounting servers.	<ul style="list-style-type: none"> Specify the primary HWTACACS accounting server: primary accounting <i>ip-address</i> [<i>port-number</i> vpn-instance <i>vpn-instance-name</i>] * Specify the secondary HWTACACS accounting server: secondary accounting <i>ip-address</i> [<i>port-number</i> vpn-instance <i>vpn-instance-name</i>] * 	<p>Configure at least one command.</p> <p>No accounting server is specified by default.</p>
4. Enable buffering of stop-accounting requests to which no responses are received.	stop-accounting-buffer enable	Optional. Enabled by default.
5. Set the maximum number of stop-accounting attempts.	retry stop-accounting <i>retry-times</i>	Optional. The default setting is 100.

Specifying the shared keys for secure HWTACACS communication

The HWTACACS client and HWTACACS server use the MD5 algorithm to authenticate packets exchanged between them and use shared keys for packet authentication and user passwords encryption. They must use the same key for the same type of communication.

To specify a shared key for secure HWTACACS communication:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter HWTACACS scheme view.	hwtacacs scheme <i>hwtacacs-scheme-name</i>	N/A
3. Specify a shared key for secure HWTACACS authentication, authorization, or accounting communication.	key { accounting authentication authorization } [cipher simple] key	No shared key is specified by default.

NOTE:

A shared key configured on the switch must be the same as that configured on the HWTACACS server.

Specifying the VPN to which the servers belong (available only on the HP 5500 EI)

After you specify a VPN for an HWTACACS scheme, all the authentication, authorization, and accounting servers specified for the scheme belong to the VPN. However, if you also specify a VPN when specifying a server for the scheme, the server belongs to the specific VPN.

To specify a VPN for an HWTACACS scheme:

Step	Command
1. Enter system view.	system-view
2. Enter HWTACACS scheme view.	hwtacacs scheme <i>hwtacacs-scheme-name</i>
3. Specify a VPN for the HWTACACS scheme.	vpn-instance <i>vpn-instance-name</i>

Setting the username format and traffic statistics units

A username is usually in the format of *userid@isp-name*, where *isp-name* represents the name of the ISP domain the user belongs to and is used by the switch to determine which users belong to which ISP domains. However, some HWTACACS servers cannot recognize usernames that contain an ISP domain name. In this case, the switch must remove the domain name of each username before sending the username. You can set the username format on the switch for this purpose.

The switch periodically sends accounting updates to HWTACACS accounting servers to report the traffic statistics of online users. For normal and accurate traffic statistics, make sure the unit for data flows and that for packets on the switch are consistent with those configured on the HWTACACS servers.

Follow these guidelines when you set the username format and the traffic statistics units for an HWTACACS scheme:

- If an HWTACACS server does not support a username that carries the domain name, configure the switch to remove the domain name before sending the username to the server.
- For level switching authentication, the **user-name-format keep-original** and **user-name-format without-domain** commands produce the same results. They make sure usernames sent to the HWTACACS server carry no ISP domain name.

To set the username format and the traffic statistics units for an HWTACACS scheme:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter HWTACACS scheme view.	hwtacacs scheme <i>hwtacacs-scheme-name</i>	N/A
3. Set the format for usernames sent to the HWTACACS servers.	user-name-format { keep-original with-domain without-domain }	Optional. By default, the ISP domain name is included in a username.
4. Specify the unit for data flows or packets sent to the HWTACACS servers.	data-flow-format { data { byte giga-byte kilo-byte mega-byte } packet { giga-packet kilo-packet mega-packet one-packet } } *	Optional. The default unit is byte for data flows and is one-packet for data packets.

Specifying a source IP address for outgoing HWTACACS packets

The source IP address of HWTACACS packets that a NAS sends must match the IP address of the NAS configured on the HWTACACS server. An HWTACACS server identifies a NAS by IP address. Upon receiving an HWTACACS packet, an HWTACACS server checks whether the source IP address of the packet is the IP address of any managed NAS. If yes, the server processes the packet. If not, the server drops the packet.

Usually, the source address of outgoing HWTACACS packets can be the IP address of the NAS's any interface that can communicate with the HWTACACS server. In some special scenarios, however, you

must change the source IP address. For example, if a Network Address Translation (NAT) device is present between the NAS and the HWTACACS server, the source IP address of outgoing HWTACACS packets must be a public IP address of the NAS. If the NAS is configured with the Virtual Router Redundancy Protocol (VRRP) for stateful failover, the source IP address of HWTACACS packets can be the virtual IP address of the VRRP group to which the uplink belongs.

You can specify the source IP address for outgoing HWTACACS packets in HWTACACS scheme view for a specific HWTACACS scheme, or in system view for all HWTACACS schemes whose servers are in a VPN or the public network.

Before sending an HWTACACS packet, a NAS selects a source IP address in this order:

1. The source IP address specified for the HWTACACS scheme.
2. The source IP address specified in system view for the VPN or public network, depending on where the HWTACACS server resides.
3. The IP address of the outbound interface specified by the route.

To specify a source IP address for all HWTACACS schemes of a VPN or the public network:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Specify a source IP address for outgoing HWTACACS packets.	hwtacacs nas-ip <i>ip-address</i> [vpn-instance <i>vpn-instance-name</i>]	By default, the IP address of the outbound interface is used as the source IP address.

To specify a source IP address for a specific HWTACACS scheme:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter HWTACACS scheme view.	hwtacacs scheme <i>hwtacacs-scheme-name</i>	N/A
3. Specify a source IP address for outgoing HWTACACS packets.	nas-ip <i>ip-address</i>	By default, the IP address of the outbound interface is used as the source IP address.

Setting timers for controlling communication with HWTACACS servers

The switch uses the following timers to control the communication with an HWTACACS server:

- **Server response timeout timer (response-timeout)**—Defines the HWTACACS request retransmission interval. After sending an HWTACACS request (authentication, authorization, or accounting request), the switch starts this timer. If the switch receives no response from the server before this timer expires, it resends the request.
- **Server quiet timer (quiet)**—Defines the duration to keep an unreachable server in blocked state. If a server is not reachable, the switch changes the server's status to blocked, starts this timer for the server, and tries to communicate with another server in active state. After this timer expires, the switch changes the status of the server back to active.
- **Real-time accounting timer (realtime-accounting)**—Defines the interval at which the switch sends real-time accounting updates to the HWTACACS accounting server for online users. To implement real-time accounting, the switch must send real-time accounting packets to the accounting server for online users periodically.

To set timers for controlling communication with HWTACACS servers:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter HWTACACS scheme view.	hwtacacs scheme <i>hwtacacs-scheme-name</i>	N/A
3. Set the HWTACACS server response timeout timer.	timer response-timeout <i>seconds</i>	Optional. The default HWTACACS server response timeout timer is 5 seconds.
4. Set the quiet timer for the primary server.	timer quiet <i>minutes</i>	Optional. The default quiet timer for the primary server is 5 minutes.
5. Set the real-time accounting interval.	timer realtime-accounting <i>minutes</i>	Optional. The default real-time accounting interval is 12 minutes.

NOTE:

Consider the performance of the NAS and the HWTACACS server when you set the real-time accounting interval. A shorter interval requires higher performance. A shorter interval requires higher performance.

Displaying and maintaining HWTACACS

Task	Command	Remarks
Display the configuration information or statistics of HWTACACS schemes .	display hwtacacs [<i>hwtacacs-server-name</i> [statistics]] [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about buffered stop-accounting requests for which no responses have been received .	display stop-accounting-buffer hwtacacs-scheme <i>hwtacacs-scheme-name</i> [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear HWTACACS statistics .	reset hwtacacs statistics { accounting all authentication authorization } [slot <i>slot-number</i>]	Available in user view
Clear buffered stop-accounting requests that get no responses.	reset stop-accounting-buffer hwtacacs-scheme <i>hwtacacs-scheme-name</i> [slot <i>slot-number</i>]	Available in user view

Configuring AAA methods for ISP domains

You configure AAA methods for an ISP domain by referencing configured AAA schemes in ISP domain view. Each ISP domain has a set of default AAA methods, which are local authentication, local authorization, and local accounting by default and can be customized. If you do not configure any AAA

methods for an ISP domain, the switch uses the system default AAA methods for authentication, authorization, and accounting of the users in the domain.

Configuration prerequisites

To use local authentication for users in an ISP domain, configure local user accounts (see "[Configuring local user attributes](#)") on the switch.

To use remote authentication, authorization, and accounting, create the required RADIUS, and HWTACACS, schemes as described in "[Configuring RADIUS schemes](#)," "[Configuring HWTACACS schemes](#)".

Creating an ISP domain

In a networking scenario with multiple ISPs, the switch may connect users of different ISPs, and users of different ISPs may have different user attributes, such as different username and password structures, different service types, and different rights. To distinguish the users of different ISPs, configure ISP domains, and configure different AAA methods and domain attributes for the ISP domains.

The switch can accommodate up to 16 ISP domains, including the system predefined ISP domain **system**. You can specify one of the ISP domains as the default domain.

On the switch, each user belongs to an ISP domain. If a user provides no ISP domain name at login, the switch considers the user belongs to the default ISP domain.

To create an ISP domain:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create an ISP domain and enter ISP domain view.	domain <i>isp-name</i>	N/A
3. Return to system view.	quit	N/A
4. Specify the default ISP domain.	domain default enable <i>isp-name</i>	Optional. By default, the default ISP domain is the system predefined ISP domain system .

NOTE:

To delete the ISP domain that is functioning as the default ISP domain, you must change it to a non-default ISP domain by using the **undo domain default enable** command.

Configuring ISP domain attributes

In an ISP domain, you can configure the following attributes for all users in the domain:

- Domain status:
By placing the ISP domain to the active or blocked state, you allow or deny network service requests from users in the domain.
- Maximum number of online users:
The switch controls the number of online users in a domain to ensure the system performance and service reliability.

- Idle cut:
This function enables the switch to check the traffic of each online user in the domain at the idle timeout interval, and to log out any user in the domain whose traffic during the idle timeout period is less than the specified minimum traffic.
- Self-service server location:
By using the information defined in this attribute, users can access the self-service server to manage their own accounts and passwords.
- Default authorization user profile:
If a user passes authentication but is authorized with no user profile, the switch authorizes the default user profile of the ISP domain to the user and restricts the user's behavior based on the profile.

To configure ISP domain attributes:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter ISP domain view.	domain <i>isp-name</i>	N/A
3. Place the ISP domain to the state of active or blocked.	state { active block }	Optional. By default, an ISP domain is in active state, and users in the domain can request network services.
4. Specify the maximum number of online users in the ISP domain.	access-limit enable <i>max-user-number</i>	Optional. No limit by default.
5. Configure the idle cut function.	idle-cut enable <i>minute</i> [<i>flow</i>]	Optional. Disabled by default. This command is effective for only LAN users and portal users.
6. Enable the self-service server location function and specify the URL of the self-service server.	self-service-url enable <i>url-string</i>	Optional. Disabled by default.
7. Specify the default authorization user profile.	authorization-attribute user-profile <i>profile-name</i>	Optional. By default, an ISP domain has no default authorization user profile.

NOTE:

- For more information about user profiles, see "[Configuring a user profile.](#)"
- A self-service RADIUS server, such as IMC, is required for the self-service server location function to work.

Configuring AAA authentication methods for an ISP domain

In AAA, authentication, authorization, and accounting are separate processes. Authentication refers to the interactive authentication process of username/password/user information during an access or service request. The authentication process does not send authorization information to a supplicant or trigger accounting.

AAA supports the following authentication methods:

- **No authentication (none)**—All users are trusted and no authentication is performed. Generally, do not use this method.
- **Local authentication (local)**—Authentication is performed by the NAS, which is configured with the user information, including the usernames, passwords, and attributes. Local authentication allows high speed and low cost, but the amount of information that can be stored is limited by the size of the storage space.
- **Remote authentication (scheme)**—The NAS cooperates with a RADIUS, or HWTACACS server to authenticate users. Remote authentication provides centralized information management, high capacity, high reliability, and support for centralized authentication service for multiple NASs. You can configure local or no authentication as the backup method, which is used when the remote server is not available. No authentication can only be configured for LAN users as the backup method of remote authentication.

You can configure AAA authentication to work alone without authorization and accounting. By default, an ISP domain uses the local authentication method.

Before configuring authentication methods, complete the following tasks:

1. For RADIUS or HWTACACS authentication, configure the RADIUS or HWTACACS scheme to be referenced first. The local and none authentication methods do not require a scheme.
2. Determine the access type or service type to be configured. With AAA, you can configure an authentication method for each access type and service type, limiting the authentication protocols that can be used for access.
3. Determine whether to configure an authentication method for all access types or service types.

Follow these guidelines when you configure AAA authentication methods for an ISP domain:

- The authentication method specified with the **authentication default** command is for all types of users and has a priority lower than that for a specific access type.
- With an authentication method that references a RADIUS scheme, AAA accepts only the authentication result from the RADIUS server. The Access-Accept message from the RADIUS server also carries the authorization information, but the authentication process ignores the information.
- If you specify the **radius-scheme radius-scheme-name local, hwtacacs-scheme hwtacacs-scheme-name local** option when you configure an authentication method, local authentication is the backup method and is used only when the remote server is not available.
- If you specify only the **local** or **none** keyword in an authentication method configuration command, the switch has no backup authentication method and performs only local authentication or does not perform any authentication.
- If the method for level switching authentication references an HWTACACS scheme, the switch uses the login username of a user for level switching authentication of the user by default. If the method for level switching authentication references a RADIUS scheme, the system uses the username configured for the corresponding privilege level on the RADIUS server for level switching authentication, rather than the login username. A username configured on the RADIUS server is in the format of **\$enab/level\$**, where *level* specifies the privilege level to which the user wants to switch. For example, if user **user1** of domain **aaa** wants to switch the privilege level to 3, the system uses **\$enab3@aaa\$** for authentication when the domain name is required and uses **\$enab3\$** for authentication when the domain name is not required.

To configure AAA authentication methods for an ISP domain:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter ISP domain view.	domain <i>isp-name</i>	N/A
3. Specify the default authentication method for all types of users.	authentication default { hwtacacs-scheme <i>hwtacacs-scheme-name</i> [local] local none radius-scheme <i>radius-scheme-name</i> [local] }	Optional. The default authentication method is local for all types of users.
4. Specify the authentication method for LAN users.	authentication lan-access { local none radius-scheme <i>radius-scheme-name</i> [local none] }	Optional. The default authentication method is used by default.
5. Specify the authentication method for login users.	authentication login { hwtacacs-scheme <i>hwtacacs-scheme-name</i> [local] local none radius-scheme <i>radius-scheme-name</i> [local] }	Optional. The default authentication method is used by default.
6. Specify the authentication method for portal users.	authentication portal { local none radius-scheme <i>radius-scheme-name</i> [local] }	Optional. The default authentication method is used by default.
7. Specify the authentication method for privilege level switching.	authentication super { hwtacacs-scheme <i>hwtacacs-scheme-name</i> radius-scheme <i>radius-scheme-name</i> }	Optional. The default authentication method is used by default.

Configuring AAA authorization methods for an ISP domain

In AAA, authorization is a separate process at the same level as authentication and accounting. Its responsibility is to send authorization requests to the specified authorization servers and to send authorization information to users after successful authorization. Authorization method configuration is optional in AAA configuration.

AAA supports the following authorization methods:

- **No authorization (none)**—The NAS performs no authorization exchange. After passing authentication, non-login users can access the network, FTP users can access the root directory of the NAS, and other login users have only the rights of Level 0 (visiting).
- **Local authorization (local)**—The NAS performs authorization according to the user attributes configured for users.
- **Remote authorization (scheme)**—The NAS cooperates with a RADIUS, or HWTACACS server to authorize users. RADIUS authorization is bound with RADIUS authentication. RADIUS authorization can work only after RADIUS authentication is successful, and the authorization information is carried in the Access-Accept message. HWTACACS authorization is separate from HWTACACS authentication, and the authorization information is carried in the authorization response after successful authentication. You can configure local authorization or no authorization as the backup method, which is used when the remote server is not available.

Before configuring authorization methods, complete the following tasks:

1. For HWTACACS authorization, configure the HWTACACS scheme to be referenced first. For RADIUS authorization, the RADIUS authorization scheme must be the same as the RADIUS authentication scheme. Otherwise, it does not take effect.

2. Determine the access type or service type to be configured. With AAA, you can configure an authorization scheme for each access type and service type, limiting the authorization protocols that can be used for access.
3. Determine whether to configure an authorization method for all access types or service types.

Follow these guidelines when you configure AAA authorization methods for an ISP domain:

- The authorization method specified with the **authorization default** command is for all types of users and has a priority lower than that for a specific access type.
- If you configure an authentication method and an authorization method that use RADIUS schemes for an ISP domain, the RADIUS scheme for authorization must be the same as that for authentication. If the RADIUS authorization configuration is invalid or RADIUS authorization fails, the RADIUS authentication also fails. Whenever RADIUS authorization fails, an error message is sent to the NAS, indicating that the server is not responding.
- If you specify the **radius-scheme** *radius-scheme-name* **local**, **hwtacacs-scheme** *hwtacacs-scheme-name* [**local** | **none**] option when you configure an authorization method, local authorization or no authorization is the backup method and is used only when the remote server is not available.
- If you specify only the **local** or **none** keyword in an authorization method configuration command, the switch has no backup authorization method and performs only local authorization or does not perform any authorization.

To configure AAA authorization methods for an ISP domain:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter ISP domain view.	domain <i>isp-name</i>	N/A
3. Specify the default authorization method for all types of users.	authorization default { hwtacacs-scheme <i>hwtacacs-scheme-name</i> [local] local none radius-scheme <i>radius-scheme-name</i> [local] }	Optional. The authorization method is local for all types of users.
4. Specify the command authorization method.	authorization command { hwtacacs-scheme <i>hwtacacs-scheme-name</i> [local none] local none }	Optional. The default authorization method is used by default.
5. Specify the authorization method for LAN users.	authorization lan-access { local none radius-scheme <i>radius-scheme-name</i> [local none] }	Optional. The default authorization method is used by default.
6. Specify the authorization method for login users.	authorization login { hwtacacs-scheme <i>hwtacacs-scheme-name</i> [local] local none radius-scheme <i>radius-scheme-name</i> [local] }	Optional. The default authorization method is used by default.
7. Specify the authorization method for portal users.	authorization portal { local none radius-scheme <i>radius-scheme-name</i> [local] }	Optional. The default authorization method is used by default.

Configuring AAA accounting methods for an ISP domain

In AAA, accounting is a separate process at the same level as authentication and authorization. This process sends accounting start/update/end requests to the specified accounting server. Accounting is optional.

AAA supports the following accounting methods:

- **No accounting (none)**—The system does not perform accounting for the users.
- **Local accounting (local)**—Local accounting is implemented on the NAS. It counts and controls the number of concurrent users who use the same local user account. It does not provide statistics for charging. The maximum number of concurrent users using the same local user account is set by the **access-limit** command in local user view.
- **Remote accounting (scheme)**—The NAS works with a RADIUS server or HWTACACS server for accounting. You can configure local or no accounting as the backup method, which is used when the remote server is not available.

By default, an ISP domain uses the local accounting method.

Before configuring accounting methods, complete the following tasks:

1. For RADIUS or HWTACACS accounting, configure the RADIUS or HWTACACS scheme to be referenced first. The local and none accounting methods do not require a scheme.
2. Determine the access type or service type to be configured. With AAA, you can configure an accounting method for each access type and service type, limiting the accounting protocols that can be used for access.
3. Determine whether to configure an accounting method for all access types or service types.

Follow these guidelines when you configure AAA accounting methods for an ISP domain:

- If you configure the **accounting optional** command, the limit on the number of local user connections is not effective.
- The accounting method specified with the **accounting default** command is for all types of users and has a priority lower than that for a specific access type.
- If you specify the **radius-scheme** *radius-scheme-name* **local** or **hwtacacs-scheme** *hwtacacs-scheme-name* **local** option when you configure an accounting method, local accounting is the backup method and is used only when the remote server is not available.
- If you specify only the **local** or **none** keyword in an accounting method configuration command, the switch has no backup accounting method and performs only local accounting or does not perform any accounting.
- Accounting is not supported for FTP services.

To configure AAA accounting methods for an ISP domain:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter ISP domain view.	domain <i>isp-name</i>	N/A

Step	Command	Remarks
3. Enable the accounting optional feature.	accounting optional	Optional. Disabled by default. With the accounting optional feature, a switch allows users to use network resources when no accounting server is available or communication with all accounting servers fails.
4. Specify the default accounting method for all types of users.	accounting default { hwtacacs-scheme <i>hwtacacs-scheme-name</i> [local] local none radius-scheme <i>radius-scheme-name</i> [local] }	Optional. The default accounting method is local for all types of users.
5. Specify the command accounting method.	accounting command hwtacacs-scheme <i>hwtacacs-scheme-name</i>	Optional. The default accounting method is used by default.
6. Specify the accounting method for LAN users.	accounting lan-access { local none radius-scheme <i>radius-scheme-name</i> [local none] }	Optional. The default accounting method is used by default.
7. Specify the accounting method for login users.	accounting login { hwtacacs-scheme <i>hwtacacs-scheme-name</i> [local] local none radius-scheme <i>radius-scheme-name</i> [local] }	Optional. The default accounting method is used by default.
8. Specify the accounting method for portal users.	accounting portal { local none radius-scheme <i>radius-scheme-name</i> [local] }	Optional. The default accounting method is used by default.

Tearing down user connections

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Tear down AAA user connections	cut connection { access-type { dot1x mac-authentication portal } all domain <i>isp-name</i> interface <i>interface-type</i> <i>interface-number</i> ip <i>ip-address</i> mac <i>mac-address</i> ucibindex <i>ucib-index</i> user-name <i>user-name</i> vlan <i>vlan-id</i> } [slot <i>slot-number</i>]	The command applies only to LAN and portal user connections.

Configuring a NAS ID-VLAN binding

The access locations of users can be identified by their access VLANs. In application scenarios where identifying the access locations of users is a must, configure NAS ID-VLAN bindings on the switch. Then, when a user gets online, the switch obtains the NAS ID by the access VLAN of the user and sends the NAS ID to the RADIUS server through the NAS-identifier attribute.

To configure a NAS ID-VLAN binding:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a NAS ID profile and enter NAS ID profile view.	aaa nas-id profile <i>profile-name</i>	You can apply a NAS ID profile to an interface enabled with portal. See " Configuring portal authentication. "
3. Configure a NAS ID-VLAN binding.	nas-id <i>nas-identifier</i> bind vlan <i>vlan-id</i>	By default, no NAS ID-VLAN binding exists.

Specifying the device ID used in stateful failover mode (available only on the HP 5500 EI)

Two switches working in stateful failover mode for portal services are uniquely identified by their device IDs. A device ID can only be 1 or 2. For more information about the stateful failover mode for portal services, see "[Configuring portal authentication.](#)"

Follow these guidelines when you specify the device ID used in stateful failover mode:

- Configuring or changing the device ID of a switch logs out all online users of the switch.
- HP recommends to save the configuration and reboot the switch after configuring or changing the device ID.
- The device ID is the symbol for stateful failover mode. Do not configure any device ID for a switch working in stand-alone mode.

To specify the device ID used in stateful failover mode:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Specify the device ID used in stateful failover mode.	nas device-id <i>device-id</i>	By default, a switch works in standalone mode and has no device ID.

Configuring a switch as a RADIUS server

RADIUS server functions configuration task list

Task	Remarks
Configuring a RADIUS user	Required
Specifying a RADIUS client	Required

Configuring a RADIUS user

This task is to create a RADIUS user and configure a set of attributes for the user on a switch that serves as the RADIUS server. The user attributes include the password, authorization attribute, expiration time,

and user description. After completing this task, the specified RADIUS user can use the username and password for RADIUS authentication on the switch.

You can use the **authorization-attribute** command to specify an authorization ACL and authorized VLAN, which is assigned by the RADIUS server to the RADIUS client (the NAS) after the RADIUS user passes authentication. The NAS then uses the assigned ACL and VLAN to control user access. If the assigned ACL does not exist on the NAS, ACL assignment fails and the NAS forcibly logs out the RADIUS user. If the assigned VLAN does not exist on the NAS, the NAS creates the VLAN and adds the RADIUS user or the access port to the VLAN.

To configure a RADIUS user:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a RADIUS user and enter RADIUS server user view.	radius-server user <i>user-name</i>	No RADIUS user exists by default.
3. Configure a password for the RADIUS user.	password [cipher simple] <i>password</i>	Optional. By default, no password is specified.
4. Configure the authorization attribute for the RADIUS user.	authorization-attribute { acl <i>acl-number</i> vlan <i>vlan-id</i> } *	Optional. Not configured by default.
5. Set the expiration time for the RADIUS user.	expiration-date <i>time</i>	Optional. By default, no expiration time is set, and the system does not check users' expiration time.
6. Configure a description for the RADIUS user.	description <i>text</i>	Optional. Not configured by default.

Specifying a RADIUS client

This task is to specify the IP address of a client to be managed by the RADIUS server and configure the shared key. The RADIUS server processes only the RADIUS packets sent from the specified clients.

To specify a RADIUS client

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Specify a RADIUS client.	radius-server client-ip <i>ip-address</i> [key [cipher simple] <i>string</i>]	No RADIUS client is specified by default.

NOTE:

- The IP address of a RADIUS client specified on the RADIUS server must be consistent with the source IP address of outgoing RADIUS packets configured on the RADIUS client.
- The shared key configured on the RADIUS server must be consistent with that configured on the RADIUS client.

Displaying and maintaining AAA

Task	Command	Remarks
Display the configuration information of ISP domains.	display domain [<i>isp-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about user connections.	display connection [access-type { dot1x mac-authentication portal } domain <i>isp-name</i> interface <i>interface-type interface-number</i> ip <i>ip-address</i> mac <i>mac-address</i> ucibindex <i>ucib-index</i> user-name <i>user-name</i> vlan <i>vlan-id</i>] [slot <i>slot-number</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

AAA configuration examples

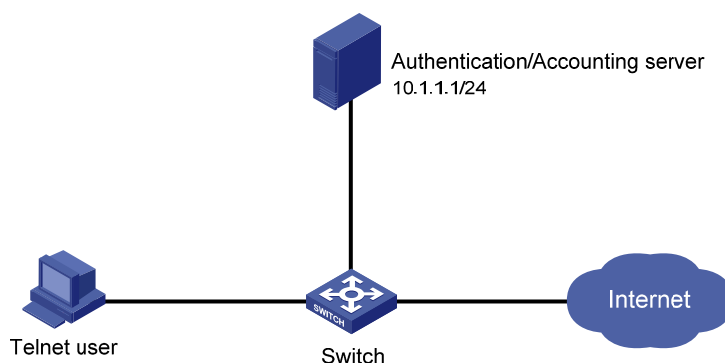
AAA for Telnet users by an HWTACACS server

Network requirements

As shown in [Figure 11](#), configure the switch to use the HWTACACS server to provide authentication, authorization, and accounting services for Telnet users.

Set the shared keys for secure communication with the HWTACACS server to **expert**. Configure the switch to remove the domain name from a username before sending the username to the HWTACACS server.

Figure 11 Network diagram



Configuration procedure

1. Configure the switch:
 - # Assign IP addresses to the interfaces. (Details not shown.)
 - # Enable the Telnet server on the switch.
 - <Switch> system-view
 - [Switch] telnet server enable
 - # Configure the switch to use AAA for Telnet users.
 - [Switch] user-interface vty 0 4

```

[Switch-ui-vty0-4] authentication-mode scheme
[Switch-ui-vty0-4] quit
# Create HWTACACS scheme hwtac.
[Switch] hwtacacs scheme hwtac
# Specify the primary authentication server.
[Switch-hwtacacs-hwtac] primary authentication 10.1.1.1 49
# Specify the primary authorization server.
[Switch-hwtacacs-hwtac] primary authorization 10.1.1.1 49
# Specify the primary accounting server.
[Switch-hwtacacs-hwtac] primary accounting 10.1.1.1 49
# Set the shared keys for secure authentication, authorization, and accounting communication to
expert.
[Switch-hwtacacs-hwtac] key authentication simple expert
[Switch-hwtacacs-hwtac] key authorization simple expert
[Switch-hwtacacs-hwtac] key accounting simple expert
# Configure the scheme to remove the domain name from a username before sending the
username to the HWTACACS server.
[Switch-hwtacacs-hwtac] user-name-format without-domain
[Switch-hwtacacs-hwtac] quit
# Configure the AAA methods for the domain.
[Switch] domain bbb
[Switch-isp-bbb] authentication login hwtacacs-scheme hwtac
[Switch-isp-bbb] authorization login hwtacacs-scheme hwtac
[Switch-isp-bbb] accounting login hwtacacs-scheme hwtac
[Switch-isp-bbb] quit

```

2. Verify the configuration:

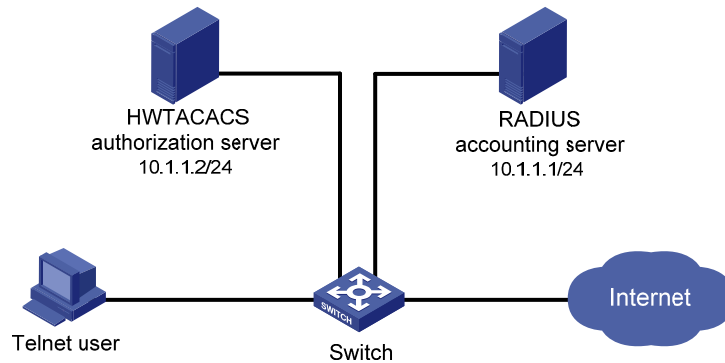
Telnet to the switch as a user and enter the correct username and password. You pass authentication and log in to the switch. Issuing the **display connection** command on the switch, you can see information about the user connection.

AAA for Telnet users by separate servers

Network requirements

As shown in [Figure 12](#), configure the switch to provide local authentication, HWTACACS authorization, and RADIUS accounting services for Telnet users. Set the shared keys for secure communication with the HWTACACS server and the RADIUS server to **expert**. Configure the switch to remove the domain name from a username before sending the username to the servers.

Figure 12 Network diagram



Configuration procedure

1. Configure the switch:

Assign IP addresses to interfaces. (Details not shown.)

Enable the Telnet server on the switch.

```
<Switch> system-view
```

```
[Switch] telnet server enable
```

Configure the switch to use AAA for Telnet users.

```
[Switch] user-interface vty 0 4
```

```
[Switch-ui-vty0-4] authentication-mode scheme
```

```
[Switch-ui-vty0-4] quit
```

Configure the HWTACACS scheme.

```
[Switch] hwtacacs scheme hwtac
```

```
[Switch-hwtacacs-hwtac] primary authorization 10.1.1.2 49
```

```
[Switch-hwtacacs-hwtac] key authorization expert
```

```
[Switch-hwtacacs-hwtac] user-name-format without-domain
```

```
[Switch-hwtacacs-hwtac] quit
```

Configure the RADIUS scheme.

```
[Switch] radius scheme rd
```

```
[Switch-radius-rd] primary accounting 10.1.1.1 1813
```

```
[Switch-radius-rd] key accounting expert
```

```
[Switch-radius-rd] server-type extended
```

```
[Switch-radius-rd] user-name-format without-domain
```

```
[Switch-radius-rd] quit
```

Create a local user named **hello**.

```
[Switch] local-user hello
```

```
[Switch-luser-hello] service-type telnet
```

```
[Switch-luser-hello] password simple hello
```

```
[Switch-luser-hello] quit
```

Configure the AAA methods for the ISP domain.

```
[Switch] domain bbb
```

```
[Switch-isp-bbb] authentication login local
```

```
[Switch-isp-bbb] authorization login hwtacacs-scheme hwtac
```

```
[Switch-isp-bbb] accounting login radius-scheme rd
```

```
[Switch-isp-bbb] quit
```

2. Verify the configuration:

Telnet to the switch as a user and enter the username **hello@bbb** and the correct password. You pass authentication and log in to the switch. Issuing the **display connection** command on the switch, you can see information about the user connection.

Authentication/authorization for SSH/Telnet users by a RADIUS server

The configuration of authentication and authorization for SSH users is similar to that for Telnet users. The following example describes the configuration for SSH users.

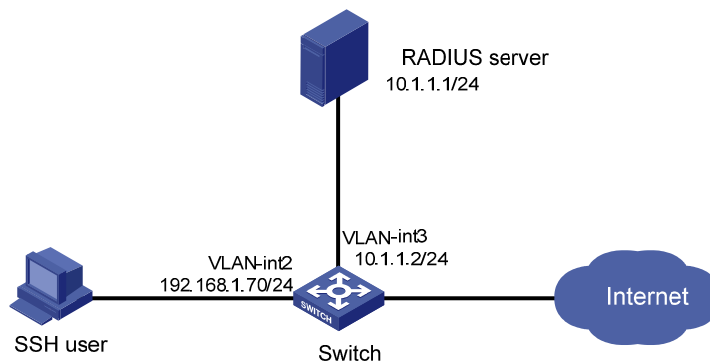
Network requirements

As shown in [Figure 13](#), configure the switch to use the RADIUS server for SSH user authentication and authorization, and to include the domain name in a username sent to the RADIUS server.

Configure IMC to act as the RADIUS server, add an account with the username **hello@bbb** on the RADIUS server, and configure the RADIUS server to assign the privilege level of 3 to the user after the user passes authentication.

Set the shared keys for secure RADIUS communication to **expert**.

Figure 13 Network diagram



Configuring the RADIUS server

This example assumes that the RADIUS server runs on IMC PLAT 5.0 (E0101) and IMC UAM 5.0 (E0101).

1. Add the switch to IMC as an access device:
 - a. Log in to IMC, click the **Service** tab, and select **User Access Manager > Access Device** from the navigation tree.
 - b. Click **Add**.
 - c. Configure the following parameters:
 - Set the shared key for secure authentication and accounting communication to **expert**.
 - Specify the ports for authentication and accounting as 1812 and 1813, respectively.
 - Select **Device Management Service** as the service type.
 - Select **HP** as the access device type.
 - Select the switch from the device list or manually add the switch with the IP address of 10.1.1.2.
 - d. Click **OK**.

NOTE:

The IP address of the access device specified here must be the same as the source IP address of the RADIUS packets sent from the switch, which is the IP address of the outbound interface by default, or otherwise the IP address specified with the **nas-ip** or **radius nas-ip** command on the switch.

Figure 14 Adding the switch to IMC as an access device

Service >> User Access Manager >> Access Device >> Add Access Device Help

Access Configuration

* Shared Key	expert	* Authentication Port	1812
* Accounting Port	1813	Service Type	Device Management S
Access Device Type	HP(A-Series)	RADIUS Accounting	Fully Supported
Service Group	Ungrouped	Access Area	--

Device List

Select Add Manually Clear All Click OK to save your change.

Total Items: 1.

Device Name	Device IP	Device Model	Delete
	10.1.1.2		X

OK Cancel

2. Add a user for device management:
 - a. Click the **User** tab, and select **Device Management User** from the navigation tree.
 - b. Click **Add**.
 - c. Configure the following parameters:
 - Enter **hello@bbb** as the username and set the password.
 - Select **SSH** as the service type.
 - Set the EXEC privilege level to 3. This value identifies the privilege level of the SSH user after login and defaults to 0.
 - Specify the IP address range of the hosts to be managed as 10.1.1.0 through 10.1.1.255.
 - d. Click **OK**.

Figure 15 Adding an account for device management

User >> Device Management User >> Add Device Management User

Add Device Management User

Basic Information of Device Management User

* Account Name ?

* User Password

* Confirm Password

Service Type ▾

EXEC Priority ?

Bound User IP List

No match found.

<input type="checkbox"/>	Start IP	End IP	Delete

IP Address List of Managed Devices

Total Items: 1.

<input type="checkbox"/>	Start IP	End IP	Delete
<input type="checkbox"/>	10.1.1.0	10.1.1.255	✘

Configuring the switch

Configure the IP address of VLAN interface 2, through which the SSH user accesses the switch.

```
<Switch> system-view
[Switch] interface vlan-interface 2
[Switch-Vlan-interface2] ip address 192.168.1.70 255.255.255.0
[Switch-Vlan-interface2] quit
```

Configure the IP address of VLAN-interface 3, through which the switch access the server.

```
[Switch] interface vlan-interface 3
[Switch-Vlan-interface3] ip address 10.1.1.2 255.255.255.0
[Switch-Vlan-interface3] quit
```

Generate RSA and DSA key pairs and enable the SSH server.

```
[Switch] public-key local create rsa
[Switch] public-key local create dsa
[Switch] ssh server enable
```

Configure the switch to use AAA for SSH users.

```
[Switch] user-interface vty 0 4
[Switch-ui-vty0-4] authentication-mode scheme
```

Configure the user interfaces to support SSH.

```
[Switch-ui-vty0-4] protocol inbound ssh
[Switch-ui-vty0-4] quit
```

Create RADIUS scheme **rad**.


```

[Switch] radius scheme rad
# Specify the primary authentication server.
[Switch-radius-rad] primary authentication 10.1.1.1 1812
# Set the shared key for secure authentication communication to expert.
[Switch-radius-rad] key authentication expert
# Configure the scheme to include the domain names in usernames to be sent to the RADIUS server.
[Switch-radius-rad] user-name-format with-domain
# Specify the service type for the RADIUS server, which must be extended when the RADIUS server runs
on IMC.
[Switch-radius-rad] server-type extended
[Switch-radius-rad] quit
# Configure the AAA methods for the domain.
[Switch] domain bbb
[Switch-isp-bbb] authentication login radius-scheme rad
[Switch-isp-bbb] authorization login radius-scheme rad
[Switch-isp-bbb] quit

```

Verifying the configuration

After you complete the configuration, the SSH user should be able to use the configured account to access the user interface of the switch and can access the demands of level 0 through level 3. .

Use the **display connection** command to view the connection information on the switch.

```

[Switch] display connection
Index=1      ,Username=hello@bbb
IP=192.168.1.58
IPv6=N/A
Total 1 connection(s) matched.

```

AAA for portal users by a RADIUS server

Network requirements

As shown in [Figure 16](#), the host automatically obtains a public network IP address through DHCP.

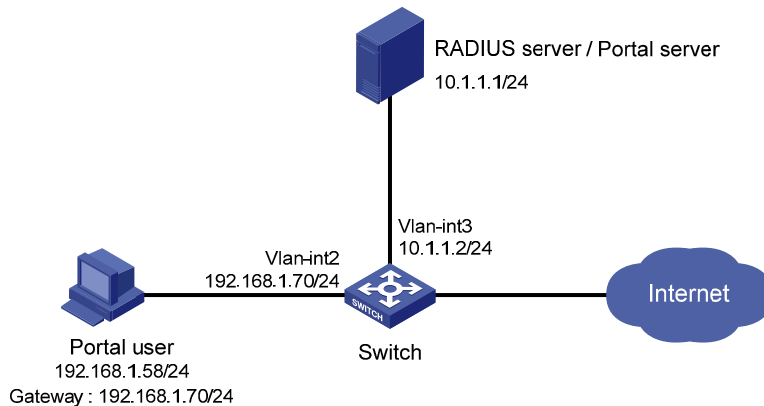
Configure the switch to:

- Use the RADIUS server for authentication, authorization, and accounting of portal users.
- Provide direct portal authentication so that the host can access only the portal server before passing portal authentication and can access the Internet after passing portal authentication.
- Keep the domain names in usernames sent to the RADIUS server.

On the RADIUS server, add a service that charges 120 dollars for up to 120 hours per month, create an account for portal users, and assign the service to the account.

Set the shared keys for secure RADIUS communication to **expert**. Set the ports for authentication/authorization and accounting to 1812 and 1813, respectively.

Figure 16 Network diagram



Configuration prerequisites

Configure IP addresses for the devices as shown in [Figure 16](#) and make sure devices can reach each other.

Configuring the RADIUS server

This example assumes that the RADIUS/portal server runs on IMC PLAT 5.0 (E0101), IMC UAM 5.0 (E0101), and IMC CAMS 5.0 (E0101).

1. Add the switch to IMC as an access device:
 - a. Log in to IMC, click the **Service** tab, and select **User Access Manager > Access Device** from the navigation tree.
 - b. Click **Add**.
 - c. Configure the following parameters:
 - Set the shared key for secure authentication and accounting communication to **expert**.
 - Specify the ports for authentication and accounting as 1812 and 1813, respectively.
 - Select **LAN Access Service** as the service type.
 - Select **HP** as the access device type.
 - Select the switch from the device list or manually add the switch whose IP address is 10.1.1.2.
 - Leave the default settings in other fields.
 - d. Click **OK**.

NOTE:

The IP address of the access device specified here must be the same as the source IP address of the RADIUS packets sent from the switch, which is the IP address of the outbound interface by default, or otherwise the IP address specified with the **nas-ip** or **radius nas-ip** command on the switch.

Figure 17 Adding the switch to IMC as an access device

Service >> User Access Manager >> Access Device >> Add Access Device Help

Access Configuration

* Shared Key	expert	* Authentication Port	1812
* Accounting Port	1813	Service Type	LAN Access Service
Access Device Type	HP(A-Series)	RADIUS Accounting	Fully Supported
Service Group	Ungrouped	Access Area	--

Device List

Select Add Manually Clear All Click OK to save your change.

Total Items: 1.

Device Name	Device IP	Device Model	Delete
	10.1.1.2		✘

OK Cancel

2. Define a charging policy:
 - a. Click the **Service** tab, and select **Accounting Manager > Charging Plans** from the navigation tree.
 - b. Click **Add**.
 - c. Configure the following parameters:
 - Enter **UserAcct** as the plan name.
 - Select **Flat rate** as the charging template.
 - In the **Basic Plan Settings** field, configure the fixed fee as 120 dollars per month.
 - In the **Service Usage Limit** field, set the **Usage Threshold** to 120 hours, allowing the user to access the Internet for up to 120 hours per month.
 - Leave the default settings in other fields.
 - d. Click **OK**.

Figure 18 Defining a charging policy

Service >> Accounting Manager >> Charging Plans >> Add Charging Plan

Charging Plan Setup

Basic Information

* Plan Name: UserAcct
 Charging Template: Flat rate
 Service Group: Ungrouped
 Description:

Basic Plan Settings

Charge Based on: time
 Billing Term: Monthly * Fixed Fee: 120 dollar

Service Usage Limit

Usage Threshold: 120 in hr

OK Cancel

3. Add a service:

- a. Click the **Service** tab, and select **User Access Manager > Service Configuration** from the navigation tree.
- b. Click **Add**.
- c. Configure the following parameters:

Enter **Portal-auth/acct** as the service name and **dm1** as the service suffix. The service suffix indicates the authentication domain for portal users. When the service suffix is configured, you must configure the switch to keep the domain names of usernames to be sent to the RADIUS server.

Enter **UserAcct** as the **Charging Plan**.

Configure other parameters as needed.

- d. Click **OK**.

Figure 19 Adding a service

Service >> User Access Manager >> Service Configuration >> Add Service Configuration

Add Service Configuration

Basic Information

* Service Name: Portal auth/acct Service Suffix: dm1
 * Service Group: Ungrouped
 Charging Plan: UserAcct
 Billing Term Start Type: Auto Start Date: Unlimited
 Adaptive consecutive deduction Charge Whole Term in Initial Term Charge by Day in Initial Term No Charge for Initial Term
 Description:
 LDAP Priority: Available ?

4. Create an account for portal users:
 - a. Click the **User** tab, and select **All Access Users** from the navigation tree.
 - b. Click **Add**.
 - c. Configure the following parameters:
 - Select the user **hello**, or add the user if it does not exist.
 - Enter **portal** as the account name and set the password.
 - Select the access service **Portal-auth/acct**.
 - Configure other parameters as needed.
 - d. Click **OK**.

Figure 20 Creating an account for portal users

User >> All Access Users >> Add Access User ? Help

Access account

Access Information

* User Name

* Account Name Fast Access User Computer User

* Password * Confirm Password

Allow User to Change Password Enable Password Strategy Modify Password at Next Login

Expiration Date

Max. Idle Time Minutes Max. Concurrent Logins

Account Type * Prepaid Money 元

Self-Service Recharge

Login Message

Access Service

<input type="checkbox"/>	Service Name	Service Suffix	Status	Charging Plan	Allocate IP
<input checked="" type="checkbox"/>	Portal-auth/acct	dm1	Available	UserAcct	

Configuring the portal server

1. Configuring the portal server:
 - a. Click the **Service** tab, and select **User Access Manager > Portal Service Management > Server** from the navigation tree.
 - b. Enter the URL address of the portal authentication main page in the format **http://ip:port/portal**, where *ip* and *port* are those configured during UAM installation. Usually, the default port 8080 is used. Leave the default settings for other parameters.
 - c. Click **OK**.

Figure 21 Portal server configuration

Service >> User Access Manager >> Portal Service Management >> Server

Portal Server Configuration

Basic Information

* Log Level: Info

* Request Timeout: 5 Seconds

* Server Heartbeat Interval: 20 Seconds

* User Heartbeat Interval: 5 Minutes

Portal Page: http://10.1.1.1:8080/portal

Advanced Information

Service Type List

Add

Total Items: 0.

Service Type ID	Service Type	Delete
-----------------	--------------	--------

OK

2. Configure an IP address group permitted for portal access:
 - a. Select **User Access Manager > Portal Service Management > IP Group** from the navigation tree.
 - b. Click **Add**.
 - c. Configure the following parameters:
 - Enter **Portal_user** as the IP group name.
 - Set the start IP address to **192.168.1.1** and the end IP address to **192.168.1.255**. The host IP address must be within this IP address group.
 - Select **Normal** as the action.
 - d. Click **OK**.

Figure 22 Adding an IP address group

Service >> User Access Manager >> Portal Service Management >> Portal IP Group Configuration >> Add IP Group

Add IP Group

* IP Group Name: Portal_user

* Start IP: 192.168.1.1

* End IP: 192.168.1.255

Service Group: Ungrouped

* Action: Normal

OK Cancel

3. Add the switch to IMC as a portal device:
 - a. Select **User Access Manager > Portal Service Management > Device** from the navigation tree to enter the portal device configuration page. Then, click **Add** to enter the page for adding a portal device, as shown in [Figure 23](#).
 - b. Click **Add**.
 - c. Configure the following parameters:
 - Enter **NAS** as the device name.
 - Enter **192.168.1.70** as the IP address of the interface on the switch that uses the portal service.
 - Enter **portal** as the key, which must be the same as that configured on the switch.
 - Set whether to enable IP address reallocation. Because direct portal authentication is used in this example, select **No** from the **Reallocate IP** list.
 - d. Click **OK**.

Figure 23 Adding a portal device

Service >> User Access Manager >> Portal Service Management >> Portal Device Configuration >> Add Device

Add Device

* Device Name	NAS	* IP Address	192.168.1.70
* Version	Portal 2.0	* Key	portal
* Listening Port	2000	* Local Challenge	No
* Authentication Retries	2	* Logout Retries	4
* Reallocate IP	No	* Support Server Heartbeat	No
* Support Server Heartbeat	No	* Support User Heartbeat	No
* Service Group	Ungrouped		
Device Description			

OK Cancel


4. Associate the portal device with the IP address group:
 - a. Click the **Port Group Information Management** icon  for the access device **NAS**.
 - b. Click **Add**.
 - c. Configure the following parameters:
 - Enter the port group name.
 - Select **Portal_user** as the IP address group. The IP address used by the user to access the network must be within this IP address group.
 - Leave the default settings in other fields.
 - d. Click **OK**.

Figure 24 Portal device list

Service >> User Access Manager >> Portal Service Management >> Device Add to My Favorites Help

Adding device "NAS" succeeded.

Query Device Information

Device Name Version

Service Group Query Reset

Device Information List

1-2 of 2. Page 1 of 1. Items per Page: 8 15 [50] 100 200

Device Name	Version	Service Group	IP Address	Port Group Information Management	Details	Modify	Delete
NAS	Portal 2.0	Ungrouped	192.168.1.70				

Figure 25 Port group configuration

Service >> User Access Manager >> Portal Service Management >> Portal Device Configuration >> Port Group Info Config >> Add Help

Port Group Info

Add Port Group Info

* Port Group Name	<input type="text" value="group"/>	* Language	<input type="text" value="Dynamic Detection"/>
* Start Port	<input type="text" value="0"/>	* End Port	<input type="text" value="zzzzz"/>
* Protocol	<input type="text" value="HTTP"/>	* Quick Authentication	<input type="text" value="No"/>
* NAT or Not	<input type="text" value="No"/>	* Error Transparent Transmission	<input type="text" value="Yes"/>
* Authentication Type	<input type="text" value="CHAP"/>	* IP Group	<input type="text" value="Portal_user"/>
* Heartbeat Interval	<input type="text" value="10"/> Minutes	* Heartbeat Timeout	<input type="text" value="30"/> Minutes
User Domain	<input type="text"/>	Port Group Description	<input type="text"/>
User Attribute Type	<input type="text"/>	Default Authentication Page	<input type="text" value="index_default.jsp"/>
Default Authentication Type	<input type="text" value="Web Identity AuthN"/>		

5. Validate the configuration:

Select **User Access Manager > Service Parameters > Validate System Configuration** from the navigation tree.

Configuring the switch

1. Configure a RADIUS scheme:

Create a RADIUS scheme named **rs1** and enter its view.

```
<Switch> system-view
[Switch] radius scheme rs1
```

Set the server type for the RADIUS scheme. When you use IMC, set the server type to **extended**.

```
[Switch-radius-rs1] server-type extended
```

Specify the primary authentication server and primary accounting server, and configure the keys for communication with the servers.

```
[Switch-radius-rs1] primary authentication 10.1.1.1
[Switch-radius-rs1] primary accounting 10.1.1.1
```



```
[Switch-radius-rs1] key authentication expert
[Switch-radius-rs1] key accounting expert
# Configure the scheme to keep the domain names in usernames to be sent to the RADIUS server.
[Switch-radius-rs1] user-name-format with-domain
[Switch-radius-rs1] quit
```

2. Configure an authentication domain:

Create an ISP domain named **dm1** and enter its view.

```
[Switch] domain dm1
```

Configure the ISP domain to use RADIUS scheme **rs1**.

```
[Switch-isp-dm1] authentication portal radius-scheme rs1
[Switch-isp-dm1] authorization portal radius-scheme rs1
[Switch-isp-dm1] accounting portal radius-scheme rs1
[Switch-isp-dm1] quit
```

Configure **dm1** as the default ISP domain for all users. Then, if a user enters a username without any ISP domain at login, the authentication and accounting methods of the default domain is used for the user.

```
[Switch] domain default enable dm1
```

3. Configure portal authentication:

Configure the portal server.

```
[Switch] portal server newpt ip 10.1.1.1 key portal port 50100 url
http://10.1.1.1:8080/portal
```

Enable portal authentication on the interface connecting the host.

```
[Switch] interface vlan-interface 2
[Switch-Vlan-interface2] portal server newpt method direct
[Switch-Vlan-interface2] quit
```

Verifying the configuration

The user can initiate portal authentication by using HP iNode client or by accessing a Web page. All initiated Web requests are redirected to the portal authentication page at <http://10.1.1.1:8080/portal>. Before passing portal authentication, the user can access only the authentication page. After passing portal authentication, the user can access the Internet.

After the user passes the portal authentication, use the following command to view the portal user information on the switch.

```
[Switch] display portal user interface vlan-interface 2
```

```
Index:19
```

```
State:ONLINE
```

```
SubState:NONE
```

```
ACL:NONE
```

```
Work-mode:stand-alone
```

```
MAC                IP                Vlan  Interface
```

```
-----
0015-e9a6-7cfe    192.168.1.58      2     Vlan-interface2
```

```
Total 1 user(s) matched, 1 listed.
```

Use the **display connection** command to view the connection information on the switch.

```
[Switch] display connection
```

```
Index=20  ,Username=portal@dm1
```

IP=192.168.1.58
IPv6=N/A
MAC=00-15-E9-A6-7C-FE

Total 1 connection(s) matched.

AAA for 802.1X users by a RADIUS server

Network requirements

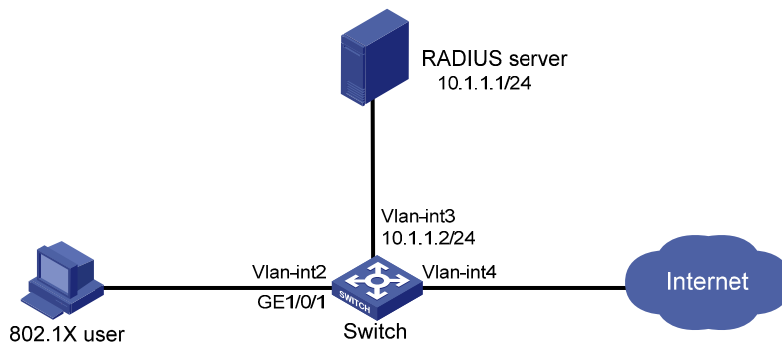
As shown in [Figure 26](#), configure the switch to:

- Use the RADIUS server for authentication, authorization, and accounting of 802.1X users.
- Use MAC-based access control on GigabitEthernet 1/0/1 to authenticate all 802.1X users on the port separately.
- Keep the domain names in usernames sent to the RADIUS server.

On the RADIUS server, add a service that charges 120 dollars for up to 120 hours per month and assigns authenticated users to VLAN 4, create an account named **dot1x@bbb** for 802.1X users, and assign the service to the account.

Set the shared keys for secure RADIUS communication to **expert**. Set the ports for authentication/authorization and accounting to 1812 and 1813, respectively.

Figure 26 Network diagram



Configuration prerequisites

Configure the interfaces and VLANs as shown in [Figure 26](#). Make sure the host can get a new IP address manually or automatically and can access resources in the authorized VLAN after passing authentication.

Configuring the RADIUS server

This example assumes that the RADIUS server runs on IMC PLAT 5.0 (E0101), IMC UAM 5.0 (E0101), and IMC CAMS 5.0 (E0101).

1. Add the switch to IMC as an access device:
 - a. Log in to IMC, click the **Service** tab, and select **User Access Manager > Access Device** from the navigation tree.
 - b. Click **Add**.
 - c. Configure the following parameters:
Set the shared key for secure authentication and accounting communication to **expert**.

Specify the ports for authentication and accounting as 1812 and 1813, respectively.

Select **LAN Access Service** as the service type.

Select **HP** as the access device type.

Select the switch from the device list or manually add the switch whose IP address is 10.1.1.2.

Leave the default settings in other fields.

- d. Click **OK**.

NOTE:

The IP address of the access device specified here must be the same as the source IP address of the RADIUS packets sent from the switch, which is the IP address of the outbound interface by default, or otherwise the IP address specified with the **nas-ip** or **radius nas-ip** command on the switch.

Figure 27 Adding the switch to IMC as an access device

Service >> User Access Manager >> Access Device >> Add Access Device Help

Access Configuration

* Shared Key	expert	* Authentication Port	1812
* Accounting Port	1813	Service Type	LAN Access Service
Access Device Type	HP(A-Series)	RADIUS Accounting	Fully Supported
Service Group	Ungrouped	Access Area	--

Device List

Select Add Manually Clear All Click OK to save your change.

Total Items: 1.

Device Name	Device IP	Device Model	Delete
	10.1.1.2		✘

OK Cancel

2. Define a charging policy:

- a. Click the **Service** tab, and select **Accounting Manager > Charging Plans** from the navigation tree.

- b. Click **Add**.

- c. Configure the following parameters:

Enter **UserAcct** as the plan name.

Select **Flat rate** as the charging template.

In the **Basic Plan Settings** field, configure to charge the fixed fee of 120 dollars per month.

In the **Service Usage Limit** field, set the **Usage Threshold** to 120 hours, allowing the user to access the Internet for up to 120 hours per month.

Leave the default settings in other fields.

- d. Click **OK**.

Figure 28 Defining a charging policy

Service >> Accounting Manager >> Charging Plans >> Add Charging Plan

Charging Plan Setup

Basic Information

* Plan Name: UserAcct

Charging Template: Flat rate

Service Group: Ungrouped

Description:

Basic Plan Settings

Charge Based on: time

Billing Term: Monthly

* Fixed Fee: 120 dollar

Service Usage Limit

Usage Threshold: 120 ? in hr

OK Cancel

3. Add a service:

- a. Click the **Service** tab, and select **User Access Manager > Service Configuration** from the navigation tree.
- b. Click **Add**.
- c. Configure the following parameters:

Enter **Dot1x auth** as the service name and **bbb** as the service suffix. The service suffix indicates the authentication domain for 802.1X users. When the service suffix is configured, you must configure the switch to keep the domain names of usernames to be sent to the RADIUS server.

Enter **UserAcct** as the **Charging Plan**.

Select **Deploy VLAN** and set the ID of the VLAN to be assigned to 4.

Configure other parameters as needed.

- d. Click **OK**.

Figure 29 Adding a service

Service >> User Access Manager >> Service Configuration >> Add Service Configuration

Add Service Configuration

Basic Information

* Service Name	<input type="text" value="Dot1x auth"/>	Service Suffix	<input type="text" value="bbb"/>
* Service Group	<input type="text" value="Ungrouped"/>		
Charging Plan	<input type="text" value="UserAcct"/>		
Billing Term Start Type	<input type="text" value="Auto"/>	Start Date	<input type="text" value="Unlimited"/>
<input type="checkbox"/> Adaptive consecutive deduction	<input checked="" type="radio"/> Charge Whole Term in Initial Term <input type="radio"/> Charge by Day in Initial Term <input type="radio"/> No Charge for Initial Term		
Description	<input type="text"/>		
LDAP Priority	<input type="text"/>	<input checked="" type="checkbox"/> Available ?	

Authorization Information

* Access Period	<input type="text" value="None"/>	Allocate IP	<input type="text" value="No"/>
Downstream Rate	<input type="text"/> Kbps	Upstream Rate	<input type="text"/> Kbps
Priority	<input type="text"/>	<input type="checkbox"/> RSA Authentication	
Certificate Authentication	<input checked="" type="radio"/> None <input type="radio"/> EAP		
Certificate Type	<input type="text" value="EAP-TLS AuthN"/>		
<input checked="" type="checkbox"/> Deploy VLAN	<input type="text" value="4"/>	<input type="checkbox"/> Deploy User Profile	<input type="text"/>
<input type="checkbox"/> Deploy User Group	<input type="text"/> ?		
<input type="checkbox"/> Deploy ACL			

4. Create an account for 802.1X users:
 - a. Click the **User** tab, and select **All Access Users** from the navigation tree.
 - b. Click **Add**.
 - c. Configure the following parameters:
 - Select the user **test**, or add the user if it does not exist.
 - Enter **dot1x** as the account name and set the password.
 - Select the access service **Dot1x auth**.
 - Configure other parameters as needed.
 - d. Click **OK**.

Figure 30 Creating an account for 802.1X users

User >> All Access Users >> Add Access User Help

Access account

Access Information

* User Name:

* Account Name: Fast Access User Computer User

* Password: * Confirm Password:

Allow User to Change Password Enable Password Strategy Modify Password at Next Login

Expiration Date:

Max. Idle Time: Minutes Max. Concurrent Logins:

Account Type: * Prepaid Money: dollar

Self-Service Recharge:

Login Message:

Access Service

	Service Name	Service Suffix	Status	Charging Plan	Allocate IP
<input checked="" type="checkbox"/>	Dot1x auth	bbb	Available	UserAcct	

Configuring the switch

1. Configure a RADIUS scheme:

Create a RADIUS scheme named **rad** and enter its view.

```
<Switch> system-view
[Switch] radius scheme rad
```

Set the server type for the RADIUS scheme. When you use IMC, set the server type to **extended**.

```
[Switch-radius-rad] server-type extended
```

Specify the primary authentication server and primary accounting server, and configure the keys for communication with the servers.

```
[Switch-radius-rad] primary authentication 10.1.1.1
[Switch-radius-rad] primary accounting 10.1.1.1
[Switch-radius-rad] key authentication expert
[Switch-radius-rad] key accounting expert
```

Configure the scheme to include the domain names in usernames to be sent to the RADIUS server.

```
[Switch-radius-rad] user-name-format with-domain
[Switch-radius-rad] quit
```

2. Configure an authentication domain:

Create an ISP domain named **bbb** and enter its view.

```
[Switch] domain bbb
```

Configure the ISP domain to use RADIUS scheme **rad**.

```
[Switch-isp-bbb] authentication lan-access radius-scheme rad
[Switch-isp-bbb] authorization lan-access radius-scheme rad
[Switch-isp-bbb] accounting lan-access radius-scheme rad
[Switch-isp-bbb] quit
```

Configure **bbb** as the default ISP domain for all users. Then, if a user enters a username without any ISP domain at login, the authentication and accounting methods of the default domain is used for the user.

```
[Switch] domain default enable bbb
```

3. Configure 802.1X authentication:

Enable 802.1X globally.

```
[Switch] dot1x
```

Enable 802.1X for port GigabitEthernet 1/0/1.

```
[Switch] interface gigabitethernet 1/0/1
```

```
[Switch-GigabitEthernet1/0/1] dot1x
```

```
[Switch-GigabitEthernet1/0/1] quit
```

Configure the access control method. (Optional. The default setting meets the requirement.)

```
[Switch] dot1x port-method macbased interface gigabitethernet 1/0/1
```

Verifying the configuration

When you use HP iNode client, no advanced authentication options are required, and the user can pass authentication after entering username **dot1x@bbb** and the correct password in the client property page.

If the 802.1X client of Windows XP is used, select the **Enable IEEE 802.1X authentication for this network** option and select **MD5-Challenge** as the EAP type on the **Authentication** tab of the network connection properties window. The user passes authentication after entering the correct username and password in the pop-up authentication page.

After the user passes authentication, the server assigns the port connecting the client to VLAN 4.

Use the **display connect** command to view the connection information on the switch.

```
[Switch] display connection
```

```
Slot: 1
```

```
Index=22 , Username=dot1x@bbb
```

```
IP=192.168.1.58
```

```
IPv6=N/A
```

```
MAC=0015-e9a6-7cfe
```

```
Total 1 connection(s) matched on slot 1.
```

```
Total 1 connection(s) matched.
```

View the information of the specified connection on the switch.

```
[Switch] display connection ucibindex 22
```

```
Slot: 1
```

```
Index=22 , Username=dot1x@bbb
```

```
IP=192.168.1.58
```

```
IPv6=N/A
```

```
MAC=0015-e9a6-7cfe
```

```
Access=8021X , AuthMethod=CHAP
```

```
Port Type=Ethernet, Port Name=GigabitEthernet1/0/1
```

```
Initial VLAN=2, Authorization VLAN=4
```

```
ACL Group=Disable
```

```
User Profile=N/A
```

```
CAR=Disable
```

```
Priority=Disable
```

```
Start=2011-04-26 19:41:12 , Current=2011-04-26 19:41:25 , Online=00h00m14s
```

Total 1 connection matched.

As the **Authorized VLAN** field in the output shows, VLAN 4 has been assigned to the user.

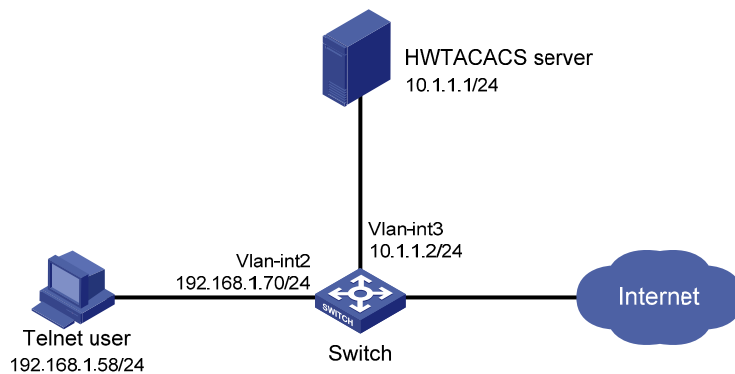
Level switching authentication for Telnet users by an HWTACACS server

Network requirements

As shown in [Figure 31](#), configure the switch to:

- Use local authentication for the Telnet user and assign the privilege level of 0 to the user after the user passes authentication.
- Use the HWTACACS server for level switching authentication of the Telnet user, and use local authentication as the backup.

Figure 31 Network diagram



Configuration considerations

1. Configure the switch to use AAA, particularly, local authentication for Telnet users:
 - Create ISP domain **bbb** and configure it to use local authentication for Telnet users.
 - Create a local user account, configure the password, and assign the user privilege level.
2. On the switch, configure the authentication method for user privilege level switching:
 - Specify to use HWTACACS authentication and, if HWTACACS authentication is not available, use local authentication for user level switching authentication.
 - Configure HWTACACS scheme **hwtac** and assign an IP address to the HWTACACS server. Set the shared keys for message exchange and specify that usernames sent to the HWTACACS server carry no domain name. Configure the domain to use the HWTACACS scheme **hwtac** for user privilege level switching authentication.
 - Configure the password for local privilege level switching authentication.
3. On the HWTACACS server, add the username and password for user privilege level switching authentication.

Configuration procedure

1. Configure the switch:

Configure the IP address of VLAN-interface 2, through which the Telnet user accesses the switch.

```
<Switch> system-view
[Switch] interface vlan-interface 2
```



```

[Switch-Vlan-interface2] ip address 192.168.1.70 255.255.255.0
[Switch-Vlan-interface2] quit
# Configure the IP address of VLAN-interface 3, through which the switch communicates with the
server.
[Switch] interface vlan-interface 3
[Switch-Vlan-interface3] ip address 10.1.1.2 255.255.255.0
[Switch-Vlan-interface3] quit
# Enable the switch to provide Telnet service.
[Switch] telnet server enable
# Configure the switch to use AAA for Telnet users.
[Switch] user-interface vty 0 4
[Switch-ui-vty0-4] authentication-mode scheme
[Switch-ui-vty0-4] quit
# Use HWTACACS authentication for user level switching authentication and, if HWTACACS
authentication is not available, use local authentication.
[Switch] super authentication-mode scheme local
# Create an HWTACACS scheme named hwtac.
[Switch] hwtacacs scheme hwtac
# Specify the IP address for the primary authentication server as 10.1.1.1 and the port for
authentication as 49.
[Switch-hwtacacs-hwtac] primary authentication 10.1.1.1 49
# Set the shared key for secure authentication communication to expert.
[Switch-hwtacacs-hwtac] key authentication simple expert
# Configure the scheme to remove the domain name from a username before sending the
username to the HWTACACS server.
[Switch-hwtacacs-hwtac] user-name-format without-domain
[Switch-hwtacacs-hwtac] quit
# Create ISP domain bbb.
[Switch] domain bbb
# Configure the ISP domain to use local authentication for Telnet users.
[Switch-isp-bbb] authentication login local
# Configure to use HWTACACS scheme hwtac for privilege level switching authentication.
[Switch-isp-bbb] authentication super hwtacacs-scheme hwtac
[Switch-isp-bbb] quit
# Create a local Telnet user named test.
[Switch] local-user test
[Switch-luser-test] service-type telnet
[Switch-luser-test] password simple aabbcc
# Configure the user level of the Telnet user to 0 after user login.
[Switch-luser-test] authorization-attribute level 0
[Switch-luser-test] quit
# Configure the password for local privilege level switching authentication to 654321.
[Switch] super password simple 654321
[Switch] quit

```

2. Configure the HWTACACS server:

NOTE:

The HWTACACS server in this example runs ACSv4.0.

Add a user named **test** on the HWTACACS server and configure advanced attributes for the user as shown in [Figure 32](#):

- Select **Max Privilege for any AAA Client** and set the privilege level to level 3. After these configurations, the user uses the password **enabpass** when switching to level 1, level 2, or level 3.
- Select **Use separate password** and specify the password as **enabpass**.

Figure 32 Configuring advanced attributes for the Telnet user

Advanced TACACS+ Settings

TACACS+ Enable Control:

- Use Group Level Setting
- No Enable Privilege
- Max Privilege for any AAA Client
Level 3

TACACS+ Enable Password

- Use CiscoSecure PAP password
- Use external database password
Windows Database

Use separate password

Password: [Redacted]

Confirm Password: [Redacted]

TACACS+ Outbound Password
(Used for SendPass and SendAuth clients such as routers)

Password: [Redacted]

Confirm Password: [Redacted]

3. Verify the configuration:

After you complete the configuration, the Telnet user should be able to telnet to the switch and use username **test@bbb** and password **aabbcc** to enter the user interface of the switch, and access all level 0 commands.

```
<Switch> telnet 192.168.1.70
Trying 192.168.1.70 ...
Press CTRL+K to abort
Connected to 192.168.1.70 ...
*****
* Copyright (c) 2004-2012 Hewlett-Packard Development Company, L.P. *
* Without the owner's prior written consent, *
* no decompiling or reverse-engineering shall be allowed. *
*****
```

Login authentication

```
Username:test@bbb
```

```
Password:
```

```
<Switch> ?
```

```
User view commands:
```

```
display  Display current system information
ping     Ping function
quit     Exit from current command view
ssh2     Establish a secure shell client connection
super    Set the current user priority level
telnet   Establish one TELNET connection
tracert  Trace route function
```

When switching to user privilege level 3, the Telnet user only needs to enter password **enabpass** as prompted.

```
<Switch> super 3
```

```
Password:
```

```
User privilege level is 3, and only those commands can be used
whose level is equal or less than this.
```

```
Privilege note: 0-VISIT, 1-MONITOR, 2-SYSTEM, 3-MANAGE
```

If the HWTACACS server is not available, the Telnet user needs to enter password **654321** as prompted for local authentication.

```
<Switch> super 3
```

```
Password: ← Enter the password for HWTACACS privilege level switch authentication
Error: Invalid configuration or no response from the authentication server.
```

```
Info: Change authentication mode to local.
```

```
Password: ← Enter the password for local privilege level switch authentication
```

```
User privilege level is 3, and only those commands can be used
whose level is equal or less than this.
```

```
Privilege note: 0-VISIT, 1-MONITOR, 2-SYSTEM, 3-MANAGE
```

RADIUS authentication and authorization for Telnet users by a switch

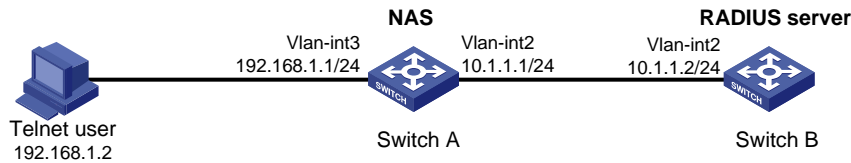
Network requirements

As shown in [Figure 33](#), configure Switch B to act as a RADIUS server to provide authentication and authorization for the Telnet user on port 1645.

Configure Switch A to use the RADIUS server for Telnet user authentication and authorization, and to remove the domain name in a username sent to the server.

Set the shared keys for secure communication between the NAS and the RADIUS server to **abc**.

Figure 33 Network diagram



Configuration procedure

1. Assign an IP address to each interface as shown in Figure 33. (Details not shown.)
2. Configure the NAS:
 - # Enable the Telnet server on Switch A.

```
<SwitchA> system-view
[SwitchA] telnet server enable
```
 - # Configure Switch A to use AAA for Telnet users.

```
[SwitchA] user-interface vty 0 4
[SwitchA-ui-vty0-4] authentication-mode scheme
[SwitchA-ui-vty0-4] quit
```
 - # Create RADIUS scheme **rad**.

```
[SwitchA] radius scheme rad
```
 - # Specify the IP address for the primary authentication server as 10.1.1.2, the port for authentication as 1645, and the shared key for secure authentication communication as **abc**.

```
[SwitchA-radius-rad] primary authentication 10.1.1.2 1645 key abc
```
 - # Configure the scheme to remove the domain name from a username before sending the username to the RADIUS server.

```
[SwitchA-radius-rad] user-name-format without-domain
```
 - # Set the source IP address for RADIUS packets as 10.1.1.1.

```
[SwitchA-radius-rad] nas-ip 10.1.1.1
[SwitchA-radius-rad] quit
```
 - # Create ISP domain **bbb**.

```
[SwitchA] domain bbb
```
 - # Specify the authentication method for Telnet users as **rad**.

```
[SwitchA-isp-bbb] authentication login radius-scheme rad
```
 - # Specify the authorization method for Telnet users as **rad**.

```
[SwitchA-isp-bbb] authorization login radius-scheme rad
```
 - # Specify the accounting method for Telnet users as **none**.

```
[SwitchA-isp-bbb] accounting login none
```
 - # Configure the RADIUS server type as **standard**. When a switch is configured to serve as a RADIUS server, the server type must be set to **standard**.

```
[SwitchA-isp-bbb] server-type standard
[SwitchA-isp-bbb] quit
```
 - # Configure **bbb** as the default ISP domain. Then, if a user enters a username without any ISP domain at login, the authentication and accounting methods of the default domain is used for the user.

```
[SwitchA] domain default enable bbb
```
3. Configure the RADIUS server:
 - # Create RADIUS user **aaa** and enter its view.

```

<SwitchB> system-view
[SwitchB] radius-server user aaa
# Configure plaintext password aabbcc for user aaa.
[SwitchB-rdsuser-aaa] password simple aabbcc
[SwitchB-rdsuser-aaa] quit
# Specify the IP address of the RADIUS client as 10.1.1.1 and the plaintext shared key as abc.
[SwitchB] radius-server client-ip 10.1.1.1 key simple abc

```

4. Verify the configuration:

After entering username **aaa@bbb** or **aaa** and password **aabbcc**, user **aaa** can telnet to Switch A. Use the **display connection** command to view the connection information on Switch A.

```

<SwitchA> display connection

Index=1      ,Username=aaa@bbb
IP=192.168.1.2
IPv6=N/A
Total 1 connection(s) matched.

```

Troubleshooting AAA

Troubleshooting RADIUS

Symptom 1

User authentication/authorization always fails.

Analysis

1. A communication failure exists between the NAS and the RADIUS server.
2. The username is not in the format of *userid@isp-name* or the ISP domain for the user authentication is not correctly configured on the NAS.
3. The user is not configured on the RADIUS server.
4. The password entered by the user is incorrect.
5. The RADIUS server and the NAS are configured with different shared key.

Solution

Check that:

1. The NAS and the RADIUS server can ping each other.
2. The username is in the *userid@isp-name* format and the ISP domain for the user authentication is correctly configured on the NAS.
3. The user is configured on the RADIUS server.
4. The correct password is entered.
5. The same shared key is configured on both the RADIUS server and the NAS.

Symptom 2

RADIUS packets cannot reach the RADIUS server.

Analysis

1. The NAS and the RADIUS server cannot communicate with each other.
2. The NAS is not configured with the IP address of the RADIUS server.
3. The UDP ports for authentication/authorization and accounting are not correct.
4. The port numbers of the RADIUS server for authentication, authorization and accounting are being used by other applications.

Solution

Check that:

1. The communication links between the NAS and the RADIUS server work well at both physical and link layers.
2. The IP address of the RADIUS server is correctly configured on the NAS.
3. UDP ports for authentication/authorization/accounting configured on the NAS are the same as those configured on the RADIUS server.
4. The port numbers of the RADIUS server for authentication, authorization and accounting are available.

Symptom 3

A user is authenticated and authorized, but accounting for the user is not normal.

Analysis

1. The accounting port number is not correct.
2. Configuration of the authentication/authorization server and the accounting server are not correct on the NAS. For example, one server is configured on the NAS to provide all the services of authentication/authorization and accounting, but in fact the services are provided by different servers.

Solution

Check that:

1. The accounting port number is correctly set.
2. The authentication/authorization server and the accounting server are correctly configured on the NAS.

Troubleshooting HWTACACS

Similar to RADIUS troubleshooting. See "[Troubleshooting RADIUS](#)."

802.1X fundamentals

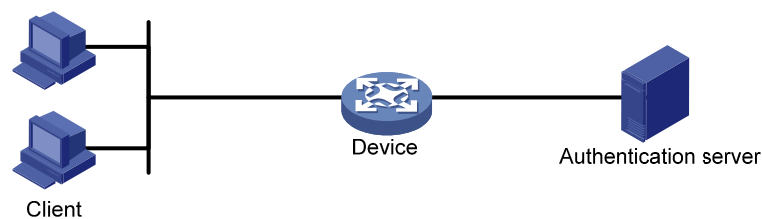
802.1X is a port-based network access control protocol initially proposed by the IEEE 802 LAN/WAN committee for securing wireless LANs (WLANs), and it has also been widely used on Ethernet networks for access control.

802.1X controls network access by authenticating the devices connected to 802.1X-enabled LAN ports.

802.1X architecture

802.1X operates in the client/server model. It comprises three entities: the client (the supplicant), the network access device (the authenticator), and the authentication server.

Figure 34 802.1X architecture



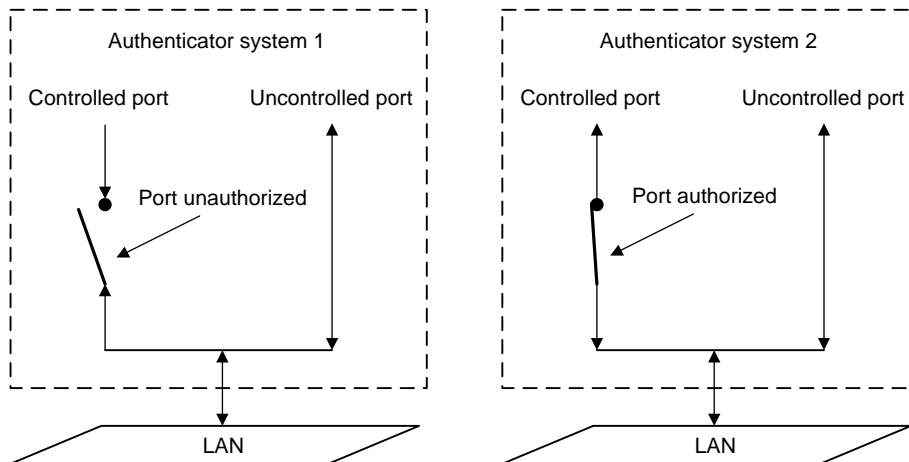
- **The client**—A user terminal seeking access to the LAN. It must have 802.1X software to authenticate to the network access device.
- **The network access device**—Authenticates the client to control access to the LAN. In a typical 802.1X environment, the network access device uses an authentication server to perform authentication.
- **The authentication server**—Provides authentication services for the network access device. It authenticates 802.1X clients by using the data sent from the network access device, and returns the authentication results for the network access device to make access decisions. The authentication server is typically a Remote Authentication Dial-in User Service (RADIUS) server. In a small LAN, you can also use the network access device as the authentication server.

Controlled/uncontrolled port and port authorization status

802.1X defines two logical ports for the network access port: controlled port and uncontrolled port. Any packet arriving at the network access port is visible to both logical ports.

- **Controlled port**—Allows incoming and outgoing traffic to pass through when it is in the authorized state, and denies incoming and outgoing traffic when it is in the unauthorized state, as shown in Figure 35. The controlled port is set in the authorized state if the client has passed authentication, and in the unauthorized state, if the client has failed authentication.
- **Uncontrolled port**—Is always open to receive and transmit EAPOL frames.

Figure 35 Authorization state of a controlled port



In the unauthorized state, a controlled port controls traffic in one of the following ways:

- Performs bidirectional traffic control to deny traffic to and from the client.
- Performs unidirectional traffic control to deny traffic from the client.

The HP devices support only unidirectional traffic control.

802.1X-related protocols

802.1X uses the Extensible Authentication Protocol (EAP) to transport authentication information for the client, the network access device, and the authentication server. EAP is an authentication framework that uses the client/server model. It supports a variety of authentication methods, including MD5-Challenge, EAP-Transport Layer Security (EAP-TLS), and Protected EAP (PEAP).

802.1X defines EAP over LAN (EAPOL) for passing EAP packets between the client and the network access device over a wired or wireless LAN. Between the network access device and the authentication server, 802.1X delivers authentication information in one of the following methods:

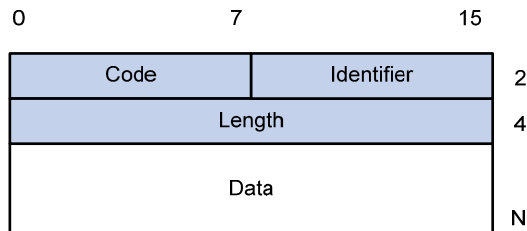
- Encapsulates EAP packets in RADIUS by using EAP over RADIUS (EAPOR), as described in "[EAP relay](#)."
- Extracts authentication information from the EAP packets and encapsulates the information in standard RADIUS packets, as described in "[EAP termination](#)."

Packet formats

EAP packet format

Figure 36 shows the EAP packet format.

Figure 36 EAP packet format

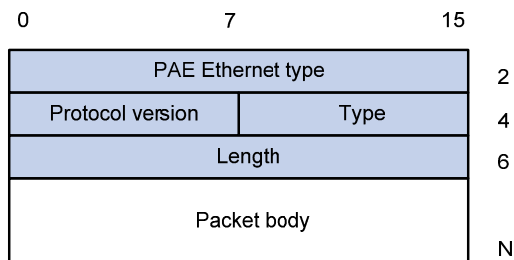


- **Code**—Type of the EAP packet. Options include Request (1), Response (2), Success (3), or Failure (4).
- **Identifier**—Used for matching Responses with Requests.
- **Length**—Length (in bytes) of the EAP packet. The EAP packet length is the sum of the Code, Identifier, Length, and Data fields.
- **Data**—Content of the EAP packet. This field appears only in a Request or Response EAP packet. The field comprises the request type (or the response type) and the type data. Type 1 (Identify) and type 4 (MD5-challenge) are two examples for the type field.

EAPOL packet format

Figure 37 shows the EAPOL packet format.

Figure 37 EAPOL packet format



- **PAE Ethernet type**—Protocol type. It takes the value 0x888E for EAPOL.
- **Protocol version**—The EAPOL protocol version used by the EAPOL packet sender.
- **Type**—Type of the EAPOL packet. Table 5 lists the types of EAPOL packets supported by HP implementation of 802.1X.

Table 5 EAPOL packet types

Value	Type	Description
0x00	EAP-Packet	The client and the network access device uses EAP-Packets to transport authentication information.
0x01	EAPOL-Start	The client sends an EAPOL-Start message to initiate 802.1X authentication to the network access device.

Value	Type	Description
0x02	EAPOL-Logoff	The client sends an EAPOL-Logoff message to tell the network access device that it is logging off.

- **Length**—Data length in bytes, or length of the Packet body. If packet type is EAPOL-Start or EAPOL-Logoff, this field is set to 0, and no Packet body field follows.
- **Packet body**—Content of the packet. When the EAPOL packet type is EAP-Packet, the Packet body field contains an EAP packet.

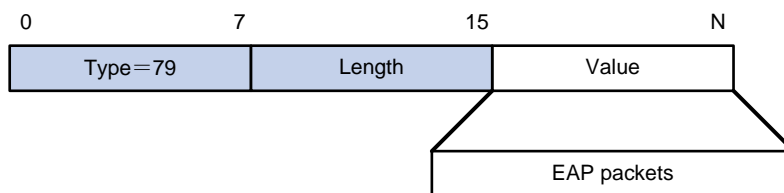
EAP over RADIUS

RADIUS adds two attributes, EAP-Message and Message-Authenticator, for supporting EAP authentication. For the RADIUS packet format, see "[Configuring AAA](#)."

EAP-Message

RADIUS encapsulates EAP packets in the EAP-Message attribute, as shown in [Figure 38](#). The Type field takes 79, and the Value field can be up to 253 bytes. If an EAP packet is longer than 253 bytes, RADIUS encapsulates it in multiple EAP-Message attributes.

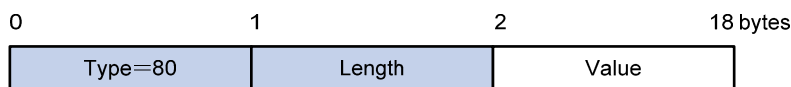
Figure 38 EAP-Message attribute format



Message-Authenticator

RADIUS includes the Message-Authenticator attribute in all packets that have an EAP-Message attribute to check their integrity. The packet receiver drops the packet if the calculated packet integrity checksum is different than the Message-Authenticator attribute value. The Message-Authenticator prevents EAP authentication packets from being tampered with during EAP authentication.

Figure 39 Message-Authenticator attribute format



Initiating 802.1X authentication

Both the 802.1X client and the access device can initiate 802.1X authentication.

802.1X client as the initiator

The client sends an EAPOL-Start packet to the access device to initiate 802.1X authentication. The destination MAC address of the packet is the IEEE 802.1X specified multicast address 01-80-C2-00-00-03 or the broadcast MAC address. If any intermediate device between the client and

the authentication server does not support the multicast address, you must use an 802.1X client, the HP iNode 802.1X client for example, that can send broadcast EAPOL-Start packets.

Access device as the initiator

The access device initiates authentication, if a client, the 802.1X client available with Windows XP for example, cannot send EAPOL-Start packets.

The access device supports the following modes:

- **Multicast trigger mode**—The access device multicasts Identity EAP-Request packets periodically (every 30 seconds by default) to initiate 802.1X authentication.
- **Unicast trigger mode**—Upon receiving a frame with the source MAC address not in the MAC address table, the access device sends an Identity EAP-Request packet out of the receiving port to the unknown MAC address. It retransmits the packet if no response has been received within a certain time interval.

802.1X authentication procedures

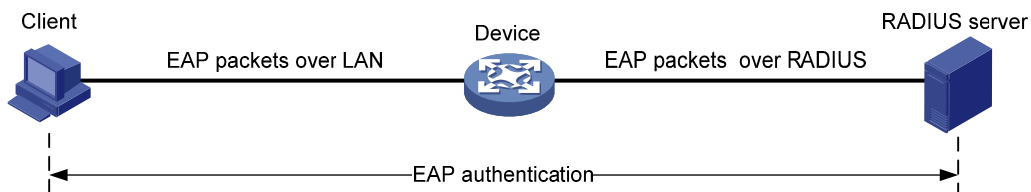
802.1X authentication has two approaches: EAP relay and EAP termination. You choose either mode depending on the support of the RADIUS server for EAP packets and EAP authentication methods.

- EAP relay mode

EAP relay is defined in IEEE 802.1X. In this mode, the network device uses EAPoR packets to send authentication information to the RADIUS server, as shown in [Figure 40](#).

In EAP relay mode, the client must use the same authentication method as the RADIUS server. On the network access device, you only need to execute the **dot1x authentication-method eap** command to enable EAP relay.

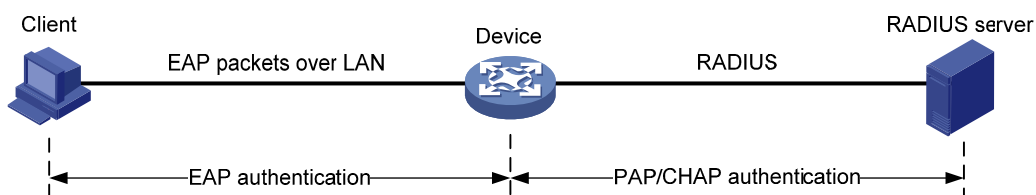
Figure 40 EAP relay



- EAP termination mode

In EAP termination mode, the network access device terminates the EAP packets received from the client, encapsulates the client authentication information in standard RADIUS packets, and uses (Password Authentication Protocol) PAP or (Password Authentication Protocol) CHAP to authenticate to the RADIUS server, as shown in [Figure 41](#).

Figure 41 EAP termination



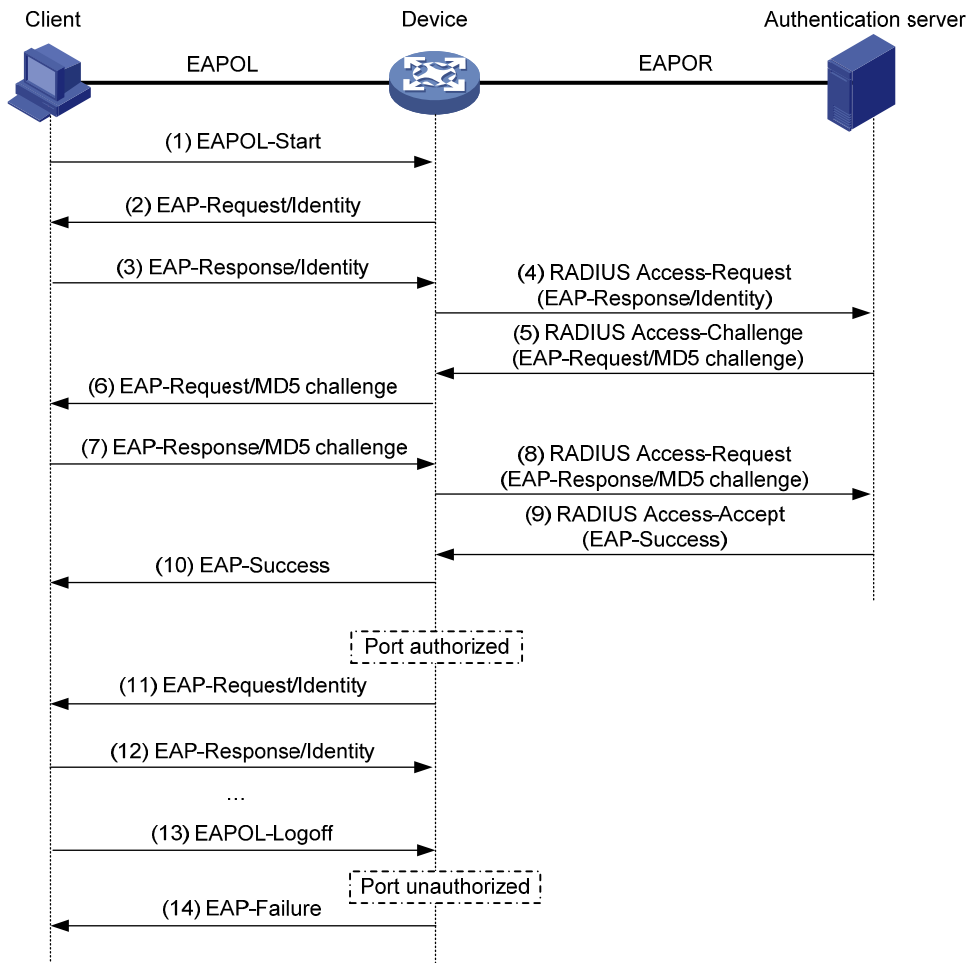
A comparison of EAP relay and EAP termination

Packet exchange method	Benefits	Limitations
EAP relay	<ul style="list-style-type: none">• Supports various EAP authentication methods.• The configuration and processing is simple on the network access device	The RADIUS server must support the EAP-Message and Message-Authenticator attributes, and the EAP authentication method used by the client.
EAP termination	Works with any RADIUS server that supports PAP or CHAP authentication.	<ul style="list-style-type: none">• Supports only MD5-Challenge EAP authentication and the "username + password" EAP authentication initiated by an HP iNode 802.1X client.• The processing is complex on the network access device.

EAP relay

Figure 42 shows the basic 802.1X authentication procedure in EAP relay mode, assuming that EAP-MD5 is used.

Figure 42 802.1X authentication procedure in EAP relay mode



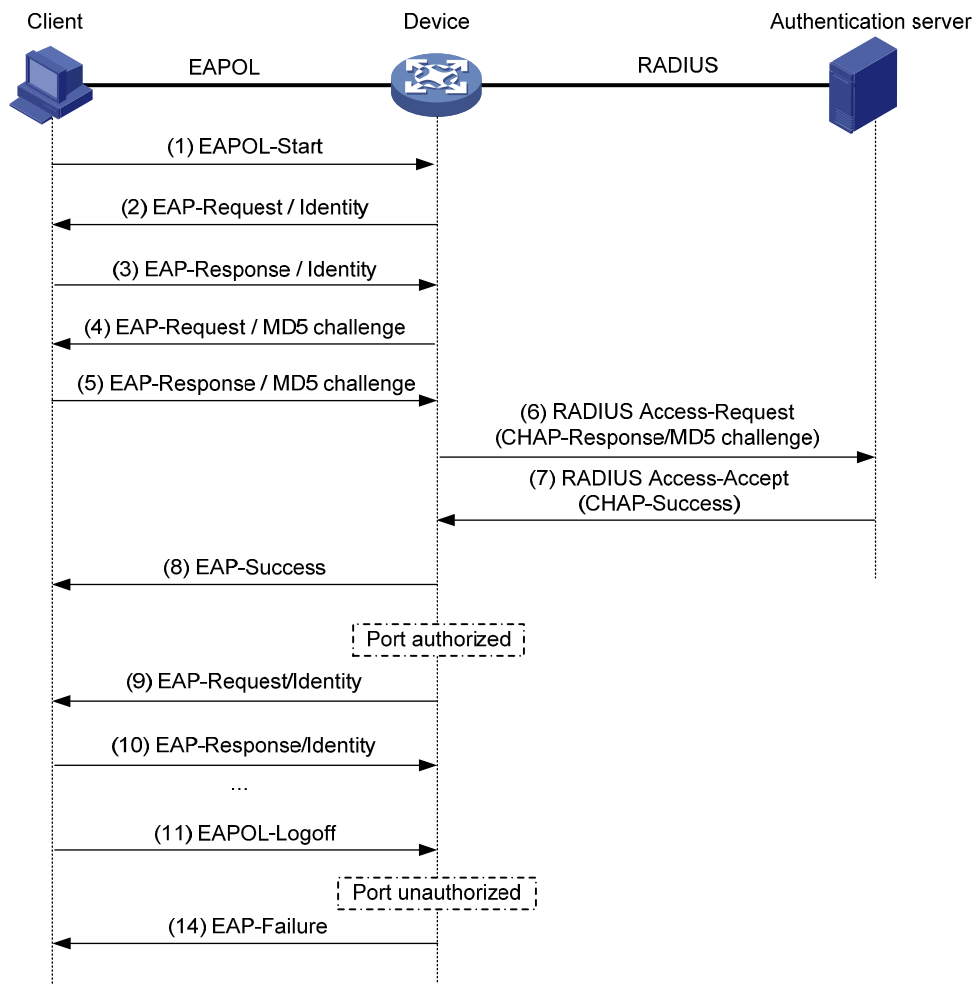
1. When a user launches the 802.1X client software and enters a registered username and password, the 802.1X client software sends an EAPOL-Start packet to the network access device.
2. The network access device responds with an Identity EAP-Request packet to ask for the client username.
3. In response to the Identity EAP-Request packet, the client sends the username in an Identity EAP-Response packet to the network access device.
4. The network access device relays the Identity EAP-Response packet in a RADIUS Access-Request packet to the authentication server.
5. The authentication server uses the identity information in the RADIUS Access-Request to search its user database. If a matching entry is found, the server uses a randomly generated challenge (EAP-Request/MD5 challenge) to encrypt the password in the entry, and sends the challenge in a RADIUS Access-Challenge packet to the network access device.
6. The network access device relays the EAP-Request/MD5 Challenge packet in a RADIUS Access-Request packet to the client.
7. The client uses the received challenge to encrypt the password, and sends the encrypted password in an EAP-Response/MD5 Challenge packet to the network access device.
8. The network access device relays the EAP-Response/MD5 Challenge packet in a RADIUS Access-Request packet to the authentication server.

9. The authentication server compares the received encrypted password with the one it generated at step 5. If the two are identical, the authentication server considers the client valid and sends a RADIUS Access-Accept packet to the network access device.
10. Upon receiving the RADIUS Access-Accept packet, the network access device sends an EAP-Success packet to the client, and sets the controlled port in the authorized state so the client can access the network.
11. After the client comes online, the network access device periodically sends handshake requests to check whether the client is still online. By default, if two consecutive handshake attempts fail, the device logs off the client.
12. Upon receiving a handshake request, the client returns a response. If the client fails to return a response after a certain number of consecutive handshake attempts (two by default), the network access device logs off the client. This handshake mechanism enables timely release of the network resources used by 802.1X users that have abnormally gone offline.
13. The client can also send an EAPOL-Logoff packet to ask the network access device for a logoff. Then
14. In response to the EAPOL-Logoff packet, the network access device changes the status of the controlled port from authorized to unauthorized and sends an EAP-Failure packet to the client.

EAP termination

Figure 43 shows the basic 802.1X authentication procedure in EAP termination mode, assuming that CHAP authentication is used.

Figure 43 802.1X authentication procedure in EAP termination mode



In EAP termination mode, it is the network access device rather than the authentication server generates an MD5 challenge for password encryption (see Step 4). The network access device then sends the MD5 challenge together with the username and encrypted password in a standard RADIUS packet to the RADIUS server.

Configuring 802.1X

This chapter describes how to configure 802.1X on an HP device.

You can also configure the port security feature to perform 802.1X. Port security combines and extends 802.1X and MAC authentication. It applies to a network that requires different authentication methods for different users on a port. Port security is beyond the scope of this chapter. It is described in "[Configuring port security](#)."

HP implementation of 802.1X

Access control methods

HP implements port-based access control as defined in the 802.1X protocol, and extends the protocol to support MAC-based access control.

- **Port-based access control**—Once an 802.1X user passes authentication on a port, any subsequent user can access the network through the port without authentication. When the authenticated user logs off, all other users are logged off.
- **MAC-based access control**—Each user is separately authenticated on a port. When a user logs off, no other online users are affected.

Using 802.1X authentication with other features

VLAN assignment

You can configure the authentication server to assign a VLAN for an 802.1X user that has passed authentication. The way that the network access device handles VLANs on an 802.1X-enabled port differs by 802.1X access control mode.

For more information about VLAN configuration and MAC-based VLAN, see *Layer 2—LAN Switching Configuration Guide*.

Access control	VLAN manipulation
Port-based	<p>Assigns the VLAN to the port as the port VLAN ID (PVID). All subsequent 802.1X users can access the port VLAN without authentication.</p> <p>When the user logs off, the previous PVID restores, and all other online users are logged off.</p>
MAC-based	<ul style="list-style-type: none">• If the port is a hybrid port with MAC-based VLAN enabled, maps the MAC address of each user to the VLAN assigned by the authentication server. The PVID of the port does not change. When a user logs off, the MAC-to-VLAN mapping for the user is removed.• If the port is an access, trunk, or MAC-based VLAN disabled hybrid port, assigns the first authenticated user's VLAN to the port as the PVID. If a different VLAN is assigned to a subsequent user, the user cannot pass the authentication. To avoid the authentication failure of subsequent users, be sure to assign the same VLAN to all 802.1X users on these ports.

With 802.1X authentication, a hybrid port is always assigned to a VLAN as an untagged member. After the assignment, do not re-configure the port as a tagged member in the VLAN.

On a periodic online user re-authentication enabled port, if a user has been online before you enable the MAC-based VLAN function, the access device does not create a MAC-to-VLAN mapping for the user unless the user passes re-authentication and the VLAN for the user has changed.

Guest VLAN

You can configure a guest VLAN on a port to accommodate users that have not performed 802.1X authentication, so they can access a limited set of network resources, such as a software server, to download anti-virus software and system patches. After a user in the guest VLAN passes 802.1X authentication, it is removed from the guest VLAN and can access authorized network resources. The way that the network access device handles VLANs on the port differs by 802.1X access control mode.

For more information about VLAN configuration and MAC-based VLAN, see *Layer 2—LAN Switching Configuration Guide*.

1. On a port that performs port-based access control

Authentication status	VLAN manipulation
No 802.1X user has performed authentication within 90 seconds after 802.1X is enabled	Assigns the 802.1X guest VLAN to the port as the PVID. All 802.1X users on this port can access only resources in the guest VLAN. If no 802.1X guest VLAN is configured, the access device does not perform any VLAN operation.
A user in the 802.1X guest VLAN fails 802.1X authentication	If an 802.1X Auth-Fail VLAN (see " Auth-Fail VLAN ") is available, assigns the Auth-Fail VLAN to the port as the PVID. All users on this port can access only resources in the Auth-Fail VLAN. If no Auth-Fail VLAN is configured, the PVID on the port is still the 802.1X guest VLAN. All users on the port are in the guest VLAN.
A user in the 802.1X guest VLAN passes 802.1X authentication	<ul style="list-style-type: none"> Assigns the VLAN specified for the user to the port as the PVID, and removes the port from the 802.1X guest VLAN. After the user logs off, the user configured PVID restores. If the authentication server assigns no VLAN, the user-configured PVID applies. The user and all subsequent 802.1X users are assigned to the user-configured port VLAN. After the user logs off, the PVID remains unchanged.

2. On a port that performs MAC-based access control

To use the 802.1X guest VLAN function on a port that performs MAC-based access control, make sure that the port is a hybrid port, and enable MAC-based VLAN on the port.

Authentication status	VLAN manipulation
A user has not passed 802.1X authentication yet	Creates a mapping between the MAC address of the user and the 802.1X guest VLAN. The user can access resources in the guest VLAN.
A user in the 802.1X guest VLAN fails 802.1X authentication	If an 802.1X Auth-Fail VLAN is available, re-maps the MAC address of the user to the Auth-Fail VLAN. The user can access only resources in the Auth-Fail VLAN. If no 802.1X Auth-Fail VLAN is configured, the user is still in the 802.1X guest VLAN.

Authentication status	VLAN manipulation
A user in the 802.1X guest VLAN passes 802.1X authentication	Re-maps the MAC address of the user to the VLAN specified for the user. If the authentication server assigns no VLAN, re-maps the MAC address of the user to the initial PVID on the port.

NOTE:

The network device assigns a hybrid port to an 802.1X guest VLAN as an untagged member.

Auth-Fail VLAN

You can configure an Auth-Fail VLAN to accommodate users that have failed 802.1X authentication because of the failure to comply with the organization security strategy, such as using a wrong password. Users in the Auth-Fail VLAN can access a limited set of network resources, such as a software server, to download anti-virus software and system patches.

The Auth-Fail VLAN does not accommodate 802.1X users that have failed authentication for authentication timeouts or network connection problems. The way that the network access device handles VLANs on the port differs by 802.1X access control mode.

For more information about VLAN configuration and MAC-based VLAN, see *Layer 2—LAN Switching Configuration Guide*.

1. On a port that performs port-based access control

Authentication status	VLAN manipulation
A user fails 802.1X authentication	Assigns the Auth-Fail VLAN to the port as the PVID. All 802.1X users on this port can access only resources in the Auth-Fail VLAN.
A user in the Auth-Fail VLAN fails 802.1X re-authentication	The Auth-Fail VLAN is still the PVID on the port, and all 802.1X users on this port are in this VLAN.
A user passes 802.1X authentication	<ul style="list-style-type: none"> • Assigns the VLAN specified for the user to the port as the PVID, and removes the port from the Auth-Fail VLAN. After the user logs off, the user-configured PVID restores. • If the authentication server assigns no VLAN, the initial PVID applies. The user and all subsequent 802.1X users are assigned to the user-configured PVID. After the user logs off, the PVID remains unchanged.

2. On a port that performs MAC-based access control

To perform the 802.1X Auth-Fail VLAN function on a port that performs MAC-based access control, you must make sure that the port is a hybrid port, and enable MAC-based VLAN on the port.

Authentication status	VLAN manipulation
A user fails 802.1X authentication	Re-maps the MAC address of the user to the Auth-Fail VLAN. The user can access only resources in the Auth-Fail VLAN.

Authentication status	VLAN manipulation
A user in the Auth-Fail VLAN fails 802.1X re-authentication	The user is still in the Auth-Fail VLAN.
A user in the Auth-Fail VLAN passes 802.1X authentication	Re-maps the MAC address of the user to the server-assigned VLAN. If the authentication server assigns no VLAN, re-maps the MAC address of the user to the initial PVID on the port.

NOTE:

The network device assigns a hybrid port to an 802.1X Auth-Fail VLAN as an untagged member.

Critical VLAN

You configure an 802.1X critical VLAN on a port to accommodate 802.1X users that fail authentication because none of the RADIUS authentication servers in their ISP domain is reachable (active). Users in the critical VLAN can access a limit set of network resources depending on your configuration.

The critical VLAN feature takes effect when 802.1X authentication is performed only through RADIUS servers. If an 802.1X user fails local authentication after RADIUS authentication, the user is not assigned to the critical VLAN. For more information about RADIUS configuration, see "[Configuring AAA](#)."

For more information about VLAN configuration and MAC-based VLAN, see *Layer 2—LAN Switching Configuration Guide*.

The way that the network access device handles VLANs on an 802.1X-enabled port differs by 802.1X access control mode.

1. On a port that performs port-based access control

Authentication status	VLAN manipulation
A user that has not been assigned to any VLAN fails 802.1X authentication because all the RADIUS servers are unreachable.	Assigns the critical VLAN to the port as the PVID. The 802.1X user and all subsequent 802.1X users on this port can access only resources in the critical VLAN.
A user in the 802.1X critical VLAN fails authentication because all the RADIUS servers are unreachable.	The critical VLAN is still the PVID of the port, and all 802.1X users on this port are in this VLAN.
A user in the 802.1X critical VLAN fails authentication for any other reason than server unreachable.	If an Auth-Fail VLAN has been configured, the PVID of the port changes to Auth-Fail VLAN ID, and all 802.1X users on this port are moved to the Auth-Fail VLAN.
A user in the critical VLAN passes 802.1X authentication.	<ul style="list-style-type: none"> • Assigns the VLAN specified for the user to the port as the PVID, and removes the port from the critical VLAN. After the user logs off, the default or user-configured PVID restores. • If the authentication server assigns no VLAN, the default or user-configured PVID applies. The user and all subsequent 802.1X users are assigned to this port VLAN. After the user logs off, this PVID remains unchanged.

Authentication status	VLAN manipulation
-----------------------	-------------------

A user in the 802.1X guest VLAN or the Auth-Fail VLAN fails authentication because all the RADIUS servers is reachable.

The PVID of the port remains unchanged. All 802.1X users on this port can access only resources in the guest VLAN or the Auth-Fail VLAN.

2. On a port that performs MAC-based access control

To perform the 802.1X critical VLAN function on a port that performs MAC-based access control, you must make sure that the port is a hybrid port, and enable MAC-based VLAN on the port.

Authentication status	VLAN manipulation
-----------------------	-------------------

A user that has not been assigned to any VLAN fails 802.1X authentication because all the RADIUS servers are unreachable.

Maps the MAC address of the user to the critical VLAN. The user can access only resources in the critical VLAN.

A user in the 802.1X critical VLAN fails authentication because all the RADIUS servers are unreachable.

The user is still in the critical VLAN.

A user in the critical VLAN fails 802.1X authentication for any other reason than server unreachable.

If an Auth-Fail VLAN has been configured, re-maps the MAC address of the user to the Auth-Fail VLAN ID.

A user in the critical VLAN passes 802.1X authentication.

Re-maps the MAC address of the user to the server-assigned VLAN.
If the authentication server assigns no VLAN, re-maps the MAC address of the user to the default or user-configured PVID on the port.

A user in the 802.1X guest VLAN or the Auth-Fail VLAN fails authentication because all the RADIUS server are unreachable.

The user remains in the 802.1X VLAN or the Auth-Fail VLAN.

A user in the MAC authentication guest VLAN fails 802.1X authentication because all the 802.1X authentication server are unreachable.

The user is removed from the MAC authentication VLAN and mapped to the 802.1X critical VLAN.

NOTE:

The network device assigns a hybrid port to an 802.1X critical VLAN as an untagged member.

Any of the following RADIUS authentication server changes in the ISP domain for 802.1X users on a port can cause the users to be removed from the critical VLAN:

- An authentication server is reconfigured, added, or removed.
- The status of any RADIUS authentication server automatically changes to active or is administratively set to active.

- The RADIUS server probing function detects that a RADIUS authentication server is reachable and sets its state to active.

You can use the **dot1x critical recovery-action reinitialize** command to configure the port to trigger 802.1X re-authentication when the port or an 802.1X user on the port is removed from the critical VLAN.

- If MAC-based access control is used, the port sends a unicast Identity EAP/Request to the 802.1X user to trigger authentication.
- If port-based access control is used, the port sends a multicast Identity EAP/Request to the 802.1X users to trigger authentication.

ACL assignment

You can specify an ACL for an 802.1X user to control its access to network resources. After the user passes 802.1X authentication, the authentication server, either the local access device or a RADIUS server, assigns the ACL to the port to filter the traffic from this user. In either case, you must configure the ACL on the access device. You can change ACL rules while the user is online.

Configuration prerequisites

- Configure an ISP domain and AAA scheme (local or RADIUS authentication) for 802.1X users.
- If RADIUS authentication is used, create user accounts on the RADIUS server.
- If local authentication is used, create local user accounts on the access device and set the service type to **lan-access**.

802.1X configuration task list

Task	Remarks
Enabling 802.1X	Required
Enabling EAP relay or EAP termination	Optional
Setting the port authorization state	Optional
Specifying an access control method	Optional
Setting the maximum number of concurrent 802.1X users on a port	Optional
Setting the maximum number of authentication request attempts	Optional
Setting the 802.1X authentication timeout timers	Optional
Configuring the online user handshake function	Optional
Configuring the authentication trigger function	Optional
Specifying a mandatory authentication domain on a port	Optional
Configuring the quiet timer	Optional
Enabling the periodic online user re-authentication function	Optional
Configuring an 802.1X guest VLAN	Optional
Configuring an Auth-Fail VLAN	Optional
Configuring an 802.1X critical VLAN	Optional
Specifying supported domain name delimiters	Optional

Enabling 802.1X

Configuration guidelines

- If the PVID of a port is a voice VLAN, the 802.1X function cannot take effect on the port. For more information about voice VLANs, see *Layer 2—LAN Switching Configuration Guide*.
- 802.1X is mutually exclusive with link aggregation and service loopback group configuration on a port.
- Do not use the BPDU drop feature on an 802.1X-enabled port. The BPDU drop feature discards 802.1X packets arrived on the port.
- On an 802.1X and MAC authentication enabled port, the EAP packet from an unknown MAC address immediately triggers 802.1X authentication, and any other type of packet from an unknown MAC address triggers MAC authentication 30 seconds after its arrival.

Configuration procedure

To enable 802.1X on a port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable 802.1X globally.	dot1x	By default, 802.1X is disabled globally.
3. Enable 802.1X on a port.	<ul style="list-style-type: none">• (Approach 1) In system view: dot1x interface <i>interface-list</i>• (Approach 2) In Ethernet interface view:<ul style="list-style-type: none">a. interface <i>interface-type</i> <i>interface-number</i>b. dot1x	Use either approach. By default, 802.1X is disabled on a port.

Enabling EAP relay or EAP termination

When you configure EAP relay or EAP termination, consider the following factors:

- The support of the RADIUS server for EAP packets
- The authentication methods supported by the 802.1X client and the RADIUS server

If the client is using only MD5-Challenge EAP authentication or the "username + password" EAP authentication initiated by an HP iNode 802.1X client, you can use both EAP termination and EAP relay. To use EAP-TL, PEAP, or any other EAP authentication methods, you must use EAP relay. When you make your decision, see "[A comparison of EAP relay and EAP termination](#)" for help.

For more information about EAP relay and EAP termination, see "[802.1X authentication procedures](#)."

To configure EAP relay or EAP termination:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure EAP relay or EAP termination.	dot1x authentication-method { chap eap pap }	Optional. By default, the network access device performs EAP termination and uses CHAP to communicate with the RADIUS server. Specify the eap keyword to enable EAP termination. Specify the chap or pap keyword to enable CHAP-enabled or PAP-enabled EAP relay.

NOTE:

If EAP relay mode is used, the **user-name-format** command configured in RADIUS scheme view does not take effect. The access device sends the authentication data from the client to the server without any modification.

Setting the port authorization state

The port authorization state determines whether the client is granted access to the network. You can control the authorization state of a port by using the **dot1x port-control** command and the following keywords:

- **authorized-force**—Places the port in the authorized state, enabling users on the port to access the network without authentication.
- **unauthorized-force**—Places the port in the unauthorized state, denying any access requests from users on the port.
- **auto**—Places the port initially in the unauthorized state to allow only EAPOL packets to pass, and after a user passes authentication, sets the port in the authorized state to allow access to the network. You can use this option in most scenarios.

You can set authorization state for one port in Ethernet interface view, or for multiple ports in system view. If different authorization state is set for a port in system view and Ethernet interface view, the one set later takes effect.

To set the authorization state of a port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the port authorization state.	<ul style="list-style-type: none"> • (Approach 1) In system view: dot1x port-control { authorized-force auto unauthorized-force } [interface <i>interface-list</i>] • (Approach 2) In Ethernet interface view: <ul style="list-style-type: none"> a. interface <i>interface-type</i> <i>interface-number</i> b. dot1x port-control { authorized-force auto unauthorized-force } 	Optional. Use either approach. By default, auto applies.

Specifying an access control method

You can specify an access control method for one port in Ethernet interface view, or for multiple ports in system view. If different access control methods are specified for a port in system view and Ethernet interface view, the one specified later takes effect.

To use both 802.1X and portal authentication on a port, you must specify MAC-based access control. For information about portal authentication, see "[Configuring portal authentication](#)."

To specify the access control method:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Specify an access control method.	<ul style="list-style-type: none">• (Approach 1) In system view: dot1x port-method { macbased portbased } [interface <i>interface-list</i>]• (Approach 2) In Ethernet interface view:<ul style="list-style-type: none">a. interface <i>interface-type interface-number</i>b. dot1x port-method { macbased portbased }	Optional. Use either approach. By default, MAC-based access control applies.

Setting the maximum number of concurrent 802.1X users on a port

You can set the maximum number of concurrent 802.1X users for ports individually in Ethernet interface view or in bulk in system view. If different settings are configured for a port in both views, the setting configured later takes effect.

To set the maximum number of concurrent 802.1X users on a port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the maximum number of concurrent 802.1X users on a port.	<ul style="list-style-type: none">• (Approach 1) In system view: dot1x max-user <i>user-number</i> [interface <i>interface-list</i>]• (Approach 2) In Ethernet interface view:<ul style="list-style-type: none">a. interface <i>interface-type interface-number</i>b. dot1x max-user <i>user-number</i> [interface <i>interface-list</i>]	Optional. Use either approach. The default maximum number of concurrent 802.1X users on a port is 256.

Setting the maximum number of authentication request attempts

The network access device retransmits an authentication request if it receives no response to the request it has sent to the client within a period of time (specified by using the **dot1x timer tx-period tx-period-value** command or the **dot1x timer supp-timeout supp-timeout-value** command). The network access device stops retransmitting the request, if it has made the maximum number of request transmission attempts but still received no response.

To set the maximum number of authentication request attempts:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the maximum number of attempts for sending an authentication request.	dot1x retry max-retry-value	Optional. The default setting is 2.

Setting the 802.1X authentication timeout timers

The network device uses the following 802.1X authentication timeout timers:

- **Client timeout timer**—Starts when the access device sends an EAP-Request/MD5 Challenge packet to a client. If no response is received when this timer expires, the access device retransmits the request to the client.
- **Server timeout timer**—Starts when the access device sends a RADIUS Access-Request packet to the authentication server. If no response is received when this timer expires, the access device retransmits the request to the server.

You can set the client timeout timer to a high value in a low-performance network, and adjust the server timeout timer to adapt to the performance of different authentication servers. In most cases, the default settings are sufficient.

To set the 802.1X authentication timeout timers:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the client timeout timer.	dot1x timer supp-timeout supp-timeout-value	Optional. The default is 30 seconds.
3. Set the server timeout timer.	dot1x timer server-timeout server-timeout-value	Optional. The default is 100 seconds.

Configuring the online user handshake function

The online user handshake function checks the connectivity status of online 802.1X users. The network access device sends handshake messages to online users at the interval specified by the **dot1x timer handshake-period** command. If no response is received from an online user after the maximum number

of handshake attempts (set by the **dot1x retry** command) has been made, the network access device sets the user in the offline state.

If iNode clients are deployed, you can also enable the online handshake security function to check for 802.1X users that use illegal client software to bypass security inspection such as proxy detection and dual network interface cards (NICs) detection. This function checks the authentication information in client handshake messages. If a user fails the authentication, the network access device logs the user off.

Configuration guidelines

Follow these guidelines when you configure the online user handshake function:

- To use the online handshake security function, make sure the online user handshake function is enabled. HP recommends that you use the iNode client software and IMC server to guarantee the normal operation of the online user handshake security function.
- If the network has 802.1X clients that cannot exchange handshake packets with the network access device, disable the online user handshake function to prevent their connections from being inappropriately torn down.

Configuration procedure

To configure the online user handshake function:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the handshake timer.	dot1x timer handshake-period <i>handshake-period-value</i>	Optional. The default is 15 seconds.
3. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Enable the online handshake function.	dot1x handshake	Optional. By default, the function is enabled.
5. Enable the online handshake security function.	dot1x handshake secure	Optional. By default, the function is disabled.

Configuring the authentication trigger function

The authentication trigger function enables the network access device to initiate 802.1X authentication when 802.1X clients cannot initiate authentication.

This function provides the following types of authentication trigger:

- **Multicast trigger**—Periodically multicasts Identity EAP-Request packets out of a port to detect 802.1X clients and trigger authentication.
- **Unicast trigger**—Enables the network device to initiate 802.1X authentication when it receives a data frame from an unknown source MAC address. The device sends a unicast Identity EAP/Request packet to the unknown source MAC address, and retransmits the packet if it has received no response within a period of time. This process continues until the maximum number of request attempts set with the **dot1x retry** command (see "[Setting the maximum number of authentication request attempts](#)") is reached.

The identity request timeout timer sets both the identity request interval for the multicast trigger and the identity request timeout interval for the unicast trigger.

Configuration guidelines

Follow these guidelines when you configure the authentication trigger function:

- Enable the multicast trigger on a port when the clients attached to the port cannot send EAPOL-Start packets to initiate 802.1X authentication.
- Enable the unicast trigger on a port if only a few 802.1X clients are attached to the port and these clients cannot initiate authentication.
- To avoid duplicate authentication packets, do not enable both triggers on a port.

Configuration procedure

To configure the authentication trigger function on a port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the username request timeout timer.	dot1x timer tx-period <i>tx-period-value</i>	Optional. The default is 30 seconds.
3. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Enable an authentication trigger.	dot1x { multicast-trigger unicast-trigger }	Required if you want to enable the unicast trigger. By default, the multicast trigger is enabled, and the unicast trigger is disabled.

Specifying a mandatory authentication domain on a port

You can place all 802.1X users in a mandatory authentication domain for authentication, authorization, and accounting on a port. No user can use an account in any other domain to access the network through the port. The implementation of a mandatory authentication domain enhances the flexibility of 802.1X access control deployment.

To specify a mandatory authentication domain for a port:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Specify a mandatory 802.1X authentication domain on the port.	dot1x mandatory-domain <i>domain-name</i>	By default, no mandatory 802.1X authentication domain is specified.

Configuring the quiet timer

The quiet timer enables the network access device to wait a period of time before it can process any authentication request from a client that has failed an 802.1X authentication.

You can set the quiet timer to a high value in a vulnerable network or a low value for quicker authentication response.

To configure the quiet timer:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable the quiet timer.	dot1x quiet-period	By default, the timer is disabled.
3. Set the quiet timer.	dot1x timer quiet-period <i>quiet-period-value</i>	Optional. The default is 60 seconds.

Enabling the periodic online user re-authentication function

Periodic online user re-authentication tracks the connection status of online users and updates the authorization attributes assigned by the server, such as the ACL, VLAN, and user profile-based QoS. The re-authentication interval is user configurable.

Configuration guidelines

- The periodic online user re-authentication timer can also be set by the authentication server in the session-timeout attribute. The server-assigned timer overrides the timer setting on the access device, and enables periodic online user re-authentication, even if the function is not configured. Support for the server assignment of re-authentication timer and the re-authentication timer configuration on the server vary with servers.
- The VLAN assignment status must be consistent before and after re-authentication. If the authentication server has assigned a VLAN before re-authentication, it must also assign a VLAN at re-authentication. If the authentication server has assigned no VLAN before re-authentication, it must not assign one at re-authentication. Violation of either rule can cause the user to be logged off. The VLANs assigned to an online user before and after re-authentication can be the same or different.
- If no critical VLAN is configured, RADIUS server unreachable can cause an online user being re-authenticated to be logged off. If a critical VLAN is configured, the user remains online and in the original VLAN.

Configuration procedure

To enable the periodic online user re-authentication function:

Step	Command	Remarks
1. Enter system view.	system-view	N/A

Step	Command	Remarks
2. Set the periodic re-authentication timer.	dot1x timer reauth-period <i>reauth-period-value</i>	Optional. The default is 3600 seconds.
3. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
4. Enable periodic online user re-authentication.	dot1x re-authenticate	By default, the function is disabled.

Configuring an 802.1X guest VLAN

Configuration guidelines

Follow these guidelines when you configure an 802.1X guest VLAN:

- You can configure only one 802.1X guest VLAN on a port. The 802.1X guest VLANs on different ports can be different.
- Assign different IDs to the voice VLAN, the port VLAN, and the 802.1X guest VLAN on a port, so the port can correctly process incoming VLAN tagged traffic.
- With 802.1X authentication, a hybrid port is always assigned to a VLAN as an untagged member. After the assignment, do not re-configure the port as a tagged member in the VLAN.
- Use [Table 6](#) when configuring multiple security features on a port.

Table 6 Relationships of the 802.1X guest VLAN and other security features

Feature	Relationship description	Reference
Super VLAN	You cannot specify a VLAN as both a super VLAN and an 802.1X guest VLAN.	See <i>Layer 2—LAN Switching Configuration Guide</i>
MAC authentication guest VLAN on a port that performs MAC-based access control	Only the 802.1X guest VLAN take effect. A user that fails MAC authentication will not be assigned to the MAC authentication guest VLAN.	See " Configuring MAC authentication "
802.1X Auth-Fail VLAN on a port that performs MAC-based access control	The 802.1X Auth-Fail VLAN has a higher priority	See " Using 802.1X authentication with other features "
Port intrusion protection on a port that performs MAC-based access control	The 802.1X guest VLAN function has higher priority than the block MAC action but lower priority than the shut down port action of the port intrusion protection feature.	See " Configuring port security "

Configuration prerequisites

- Create the VLAN to be specified as the 802.1X guest VLAN.
- If the 802.1X-enabled port performs port-based access control, enable 802.1X multicast trigger (**dot1x multicast-trigger**).

- If the 802.1X-enabled port performs MAC-based access control, configure the port as a hybrid port, enable MAC-based VLAN on the port, and assign the port to the 802.1X guest VLAN as an untagged member. For more information about the MAC-based VLAN function, see *Layer 2—LAN Switching Configuration Guide*.

Configuration procedure

To configure an 802.1X guest VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure an 802.1X guest VLAN for one or more ports.	<ul style="list-style-type: none"> • (Approach 1) In system view: dot1x guest-vlan <i>guest-vlan-id</i> [interface <i>interface-list</i>] • (Approach 2) In Ethernet interface view: <ol style="list-style-type: none"> interface <i>interface-type interface-number</i> dot1x guest-vlan <i>guest-vlan-id</i> 	<p>Use either approach.</p> <p>By default, no 802.1X guest VLAN is configured on any port.</p>

Configuring an Auth-Fail VLAN

Configuration guidelines

Follow these guidelines when configuring an 802.1X Auth-Fail VLAN:

- Assign different IDs to the voice VLAN, the port VLAN, and the 802.1X Auth-Fail VLAN on a port, so the port can correctly process VLAN tagged incoming traffic.
- You can configure only one 802.1X Auth-Fail VLAN on a port. The 802.1X Auth-Fail VLANs on different ports can be different.
- Use [Table 7](#) when configuring multiple security features on a port.

Table 7 Relationships of the 802.1X Auth-Fail VLAN with other features

Feature	Relationship description	Reference
Super VLAN	You cannot specify a VLAN as both a super VLAN and an 802.1X Auth-Fail VLAN.	See <i>Layer 2—LAN Switching Configuration Guide</i>
MAC authentication guest VLAN on a port that performs MAC-based access control	The 802.1X Auth-Fail VLAN has a high priority.	See " Configuring MAC authentication "
Port intrusion protection on a port that performs MAC-based access control	The 802.1X Auth-Fail VLAN function has higher priority than the block MAC action but lower priority than the shut down port action of the port intrusion protection feature.	See " Configuring port security "

Configuration prerequisites

- Create the VLAN to be specified as the 802.1X Auth-Fail VLAN.
- If the 802.1X-enabled port performs port-based access control, enable 802.1X multicast trigger (**dot1x multicast-trigger**).
- If the 802.1X-enabled port performs MAC-based access control, configure the port as a hybrid port, enable MAC-based VLAN on the port, and assign the port to the Auth-Fail VLAN as an untagged member. For more information about the MAC-based VLAN function, see *Layer 2—LAN Switching Configuration Guide*.

Configuration procedure

To configure an Auth-Fail VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the Auth-Fail VLAN on the port.	dot1x auth-fail vlan <i>authfail-vlan-id</i>	By default, no Auth-Fail VLAN is configured.

Configuring an 802.1X critical VLAN

Configuration guidelines

- Assign different IDs to the voice VLAN, the port VLAN, and the 802.1X critical VLAN on a port, so the port can correctly process VLAN tagged incoming traffic.
- You can configure only one 802.1X critical VLAN on a port. The 802.1X critical VLANs on different ports can be different.
- You cannot specify a VLAN as both a super VLAN and an 802.1X critical VLAN. For information about super VLANs, see *Layer 2—LAN Switching Configuration Guide*.

Configuration prerequisites

- Create the VLAN to be specified as a critical VLAN.
- If the 802.1X-enabled port performs port-based access control, enable 802.1X multicast trigger (**dot1x multicast-trigger**).
- If the 802.1X-enabled port performs MAC-based access control, configure the port as a hybrid port, enable MAC-based VLAN on the port, and assign the port to the Auth-Fail VLAN as an untagged member. For more information about the MAC-based VLAN function, see *Layer 2—LAN Switching Configuration Guide*.

Configuration procedure

To configure an 802.1X critical VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure an 802.1X critical VLAN on the port.	dot1x critical vlan <i>vlan-id</i>	By default, no critical VLAN is configured.
4. Configure the port to trigger 802.1X authentication on detection of a reachable authentication server for users in the critical VLAN.	dot1x critical recovery-action reinitialize	Optional. By default, when a reachable RADIUS server is detected, the system removes the port or 802.1X users from the critical VLAN without triggering authentication.

Specifying supported domain name delimiters

By default, the access device supports the at sign (@) as the delimiter. You can also configure the access device to accommodate 802.1X users that use other domain name delimiters.

The configurable delimiters include the at sign (@), back slash (\), and forward slash (/).

If an 802.1X username string contains multiple configured delimiters, the leftmost delimiter is the domain name delimiter. For example, if you configure @, /, and \ as delimiters, the domain name delimiter for the username string 123/22\@abc is the forward slash (/).

If a username string contains none of the delimiters, the access device authenticates the user in the mandatory or default ISP domain. The access selects a domain delimiter from the delimiter set in this order: @, /, and \.

Follow the steps to specify a set of domain name delimiters:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Specify a set of domain name delimiters for 802.1X users.	dot1x domain-delimiter <i>string</i>	Optional. By default, only the at sign (@) delimiter is supported.

NOTE:

If you configure the access device to include the domain name in the username sent to the RADIUS server, make sure the domain delimiter in the username can be recognized by the RADIUS server. For username format configuration, see the **user-name-format** command in *Security Command Reference*.

Displaying and maintaining 802.1X

Task	Command	Remarks
Display 802.1X session information, statistics, or configuration information of specified or all ports.	<code>display dot1x [sessions statistics] [interface interface-list] [{ begin exclude include } regular-expression]</code>	Available in any view
Clear 802.1X statistics.	<code>reset dot1x statistics [interface interface-list]</code>	Available in user view

802.1X authentication configuration example

Network requirements

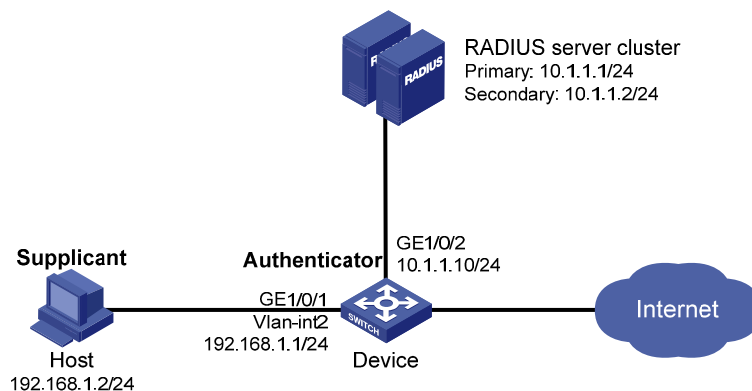
As shown in [Figure 44](#), the access device performs 802.1X authentication for users that connect to port GigabitEthernet 1/0/1. Implement MAC-based access control on the port, so the logoff of one user does not affect other online 802.1X users.

Use RADIUS servers to perform authentication, authorization, and accounting for the 802.1X users. If RADIUS authentication fails, perform local authentication on the access device. If RADIUS accounting fails, the access device logs the user off.

Configure the host at 10.1.1.1 as the primary authentication and accounting servers, and the host at 10.1.1.2 as the secondary authentication and accounting servers. Assign all users to the ISP domain **aabcc.net**, which accommodates up to 30 users.

Configure the shared key as **name** for packets between the access device and the authentication server, and the shared key as **money** for packets between the access device and the accounting server.

Figure 44 Network diagram



Configuration procedure

1. Configure the 802.1X client. If HP iNode is used, do not select the **Carry version info** option in the client configuration. (Details not shown.)
2. Configure the RADIUS servers and add user accounts for the 802.1X users. For information about the RADIUS commands used on the access device in this example, see *Security Command Reference*. (Details not shown.)
3. Assign an IP address to each interface on the access device. (Details not shown.)
4. Configure user accounts for the 802.1X users on the access device:

Add a local user with the username **localuser**, and password **localpass** in plaintext. (Make sure the username and password are the same as those configured on the RADIUS server.)

```
<Device> system-view
[Device] local-user localuser
[Device-luser-localuser] service-type lan-access
[Device-luser-localuser] password simple localpass
# Configure the idle cut function to log off any online user that has been idled for 20 minutes.
[Device-luser-localuser] authorization-attribute idle-cut 20
[Device-luser-localuser] quit
```

5. Configure a RADIUS scheme:

Create the RADIUS scheme **radius1** and enter its view.

```
[Device] radius scheme radius1
# Specify the IP addresses of the primary authentication and accounting RADIUS servers.
[Device-radius-radius1] primary authentication 10.1.1.1
[Device-radius-radius1] primary accounting 10.1.1.1
# Configure the IP addresses of the secondary authentication and accounting RADIUS servers.
[Device-radius-radius1] secondary authentication 10.1.1.2
[Device-radius-radius1] secondary accounting 10.1.1.2
# Specify the shared key between the access device and the authentication server.
[Device-radius-radius1] key authentication name
# Specify the shared key between the access device and the accounting server.
[Device-radius-radius1] key accounting money
# Exclude the ISP domain name from the username sent to the RADIUS servers.
[Device-radius-radius1] user-name-format without-domain
[Device-radius-radius1] quit
```

NOTE:

The access device must use the same username format as the RADIUS server. If the RADIUS server includes the ISP domain name in the username, so must the access device.

6. Configure the ISP domain:

Create the ISP domain **aabbcc.net** and enter its view.

```
[Device] domain aabbcc.net
# Apply the RADIUS scheme radius1 to the ISP domain, and specify local authentication as the secondary authentication method.
[Device-isp-aabbcc.net] authentication lan-access radius-scheme radius1 local
[Device-isp-aabbcc.net] authorization lan-access radius-scheme radius1 local
[Device-isp-aabbcc.net] accounting lan-access radius-scheme radius1 local
# Set the maximum number of concurrent users in the domain to 30.
[Device-isp-aabbcc.net] access-limit enable 30
# Configure the idle cut function to log off any online domain user that has been idle for 20 minutes.
[Device-isp-aabbcc.net] idle-cut enable 20
[Device-isp-aabbcc.net] quit
# Specify aabbcc.net as the default ISP domain. If a user does not provide any ISP domain name, it is assigned to the default ISP domain.
```

```
[Device] domain default enable aabbcc.net
```

7. Configure 802.1X:

```
# Enable 802.1X globally.
```

```
[Device] dot1x
```

```
# Enable 802.1X on port GigabitEthernet 1/0/1.
```

```
[Device] interface gigabitethernet 1/0/1
```

```
[Device-GigabitEthernet1/0/1] dot1x
```

```
[Device-GigabitEthernet1/0/1] quit
```

```
# Enable MAC-based access control on the port. (Optional. MAC-based access control is the default setting.)
```

```
[Device] dot1x port-method macbased interface gigabitethernet 1/0/1
```

Verifying the configuration

Use the **display dot1x interface gigabitethernet 1/0/1** command to verify the 802.1X configuration. After an 802.1X user passes RADIUS authentication, you can use the **display connection** command to view the user connection information. If the user fails RADIUS authentication, local authentication is performed.

802.1X with guest VLAN and VLAN assignment configuration example

Network requirements

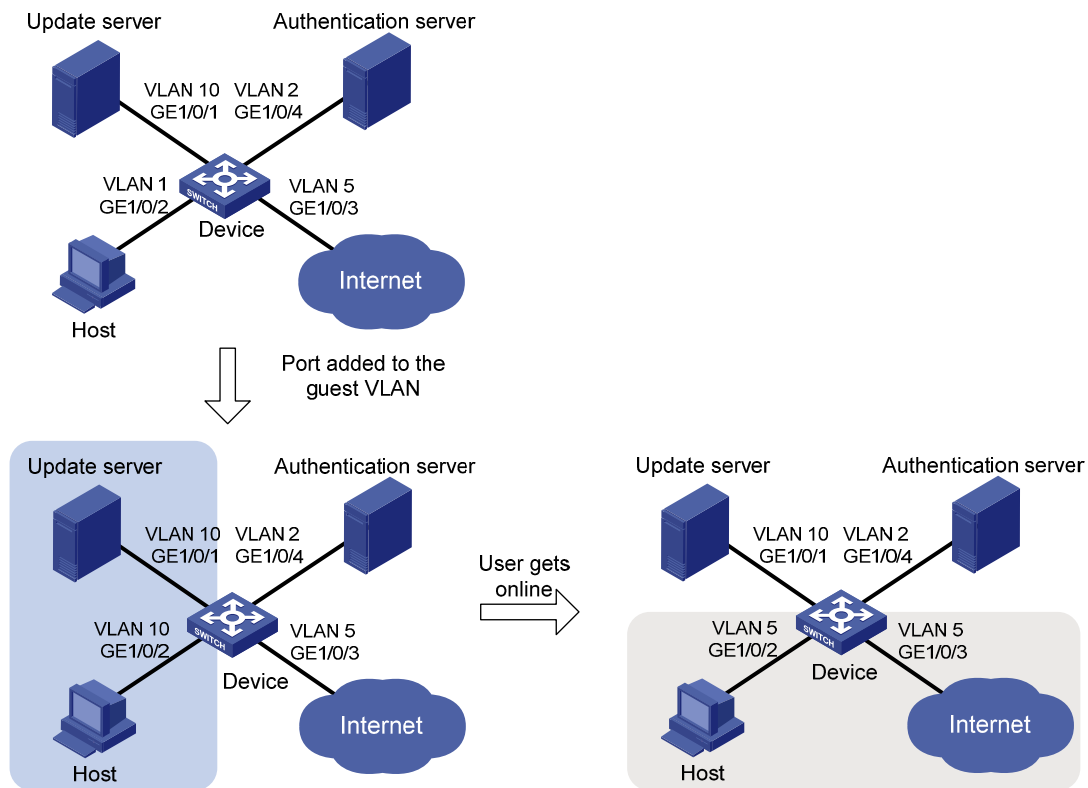
As shown in [Figure 45](#):

- A host is connected to port GigabitEthernet 1/0/2 of the device and must pass 802.1X authentication to access the Internet. GigabitEthernet 1/0/2 is in VLAN 1.
- GigabitEthernet 1/0/2 implements port-based access control.
- GigabitEthernet 1/0/3 is in VLAN 5 and is for accessing the Internet.
- The authentication server runs RADIUS and is in VLAN 2.
- The update server in VLAN 10 is for client software download and upgrade.

If no user performs 802.1X authentication on GigabitEthernet 1/0/2 within a period of time, the device adds GigabitEthernet 1/0/2 to its guest VLAN, VLAN 10. The host and the update server are both in VLAN 10 and the host can access the update server and download the 802.1X client software.

After the host passes 802.1X authentication, the network access device assigns the host to VLAN 5 where GigabitEthernet 1/0/3 is. The host can access the Internet.

Figure 45 Network diagram



Configuration procedure

The following configuration procedure covers most AAA/RADIUS configuration commands on the device. The configuration on the 802.1X client and RADIUS server are not shown. For more information about AAA/RADIUS configuration commands, see *Security Command Reference*.

1. Make sure the 802.1X client can update its IP address after the access port is assigned to the guest VLAN or a server-assigned VLAN. (Details not shown.)
2. Configure the RADIUS server to provide authentication, authorization, and accounting services. Configure user accounts and server-assigned VLAN, VLAN 5 in this example. (Details not shown.)
3. Create VLANs, and assign ports to the VLANs.

```
<Device> system-view
[Device] vlan 1
[Device-vlan1] port gigabitethernet 1/0/2
[Device-vlan1] quit
[Device] vlan 10
[Device-vlan10] port gigabitethernet 1/0/1
[Device-vlan10] quit
[Device] vlan 2
[Device-vlan2] port gigabitethernet 1/0/4
[Device-vlan2] quit
[Device] vlan 5
[Device-vlan5] port gigabitethernet 1/0/3
[Device-vlan5] quit
```

4. Configure a RADIUS scheme:

```
# Configure RADIUS scheme 2000 and enter its view.
<Device> system-view
[Device] radius scheme 2000

# Specify primary and secondary authentication and accounting servers. Set the shared key to abc
for authentication and accounting packets.
[Device-radius-2000] primary authentication 10.11.1.1 1812
[Device-radius-2000] primary accounting 10.11.1.1 1813
[Device-radius-2000] key authentication abc
[Device-radius-2000] key accounting abc

# Exclude the ISP domain name from the username sent to the RADIUS server.
[Device-radius-2000] user-name-format without-domain
[Device-radius-2000] quit
```

5. Configure an ISP domain:

```
# Create ISP domain bbb and enter its view.
[Device] domain bbb

# Apply RADIUS scheme 2000 to the ISP domain for authentication, authorization, and
accounting.
[Device-isp-bbb] authentication lan-access radius-scheme 2000
[Device-isp-bbb] authorization lan-access radius-scheme 2000
[Device-isp-bbb] accounting lan-access radius-scheme 2000
[Device-isp-bbb] quit
```

6. Configure 802.1X:

```
# Enable 802.1X globally.
[Device] dot1x

# Enable 802.1X for port GigabitEthernet 1/0/2.
[Device] interface gigabitethernet 1/0/2
[Device-GigabitEthernet1/0/2] dot1x

# Implement port-based access control on the port.
[Device-GigabitEthernet1/0/2] dot1x port-method portbased

# Set the port authorization mode to auto. This step is optional. By default, the port is in auto mode.
[Device-GigabitEthernet1/0/2] dot1x port-control auto
[Device-GigabitEthernet1/0/2] quit

# Set VLAN 10 as the 802.1X guest VLAN for port GigabitEthernet 1/0/2.
[Device] dot1x guest-vlan 10 interface gigabitethernet 1/0/2
```

Verifying the configuration

Use the **display dot1x interface gigabitethernet 1/0/2** command to verify the 802.1X guest VLAN configuration on GigabitEthernet 1/0/2. If no user passes authentication on the port within a specific period of time, use the **display vlan 10** command to verify whether GigabitEthernet 1/0/2 is assigned to VLAN 10.

After a user passes authentication, you can use the **display interface gigabitethernet 1/0/2** command to verify that port GigabitEthernet 1/0/2 has been added to VLAN 5.

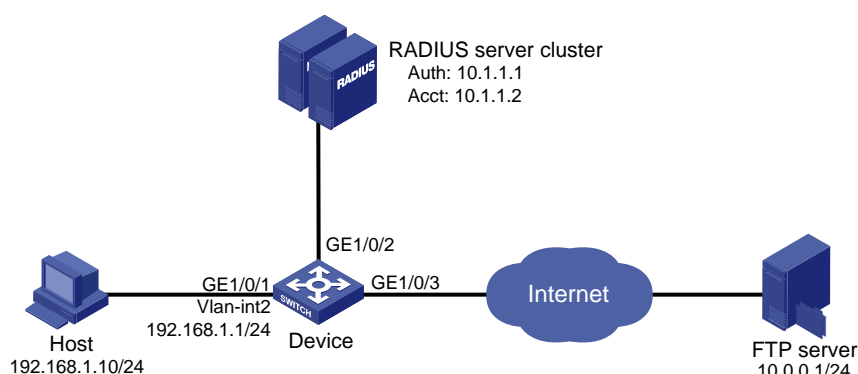
802.1X with ACL assignment configuration example

Network requirements

As shown in [Figure 46](#), the host at 192.168.1.10 connects to port GigabitEthernet 1/0/1 of the network access device.

Perform 802.1X authentication on the port. Use the RADIUS server at 10.1.1.1 as the authentication and authorization server and the RADIUS server at 10.1.1.2 as the accounting server. Assign an ACL to GigabitEthernet 1/0/1 to deny the access of 802.1X users to the FTP server at 10.0.0.1/24 on weekdays during business hours from 8:00 to 18:00.

Figure 46 Network diagram



Configuration procedure

The following configuration procedure provides the major AAA and RADIUS configuration on the access device. The configuration procedures on the 802.1X client and RADIUS server are beyond the scope of this configuration example. For information about AAA and RADIUS configuration commands, see *Security Command Reference*.

1. Configure 802.1X client. Make sure the client is able to update its IP address after the access port is assigned to the 802.1X guest VLAN or a server-assigned VLAN. (Details not shown.)
2. Configure the RADIUS servers, user accounts, and authorization ACL, ACL 3000 in this example. (Details not shown.)
3. Configure the access device:

Assign IP addresses to interfaces. (Details not shown.)

Configure the RADIUS scheme.

```
<Device> system-view
[Device] radius scheme 2000
[Device-radius-2000] primary authentication 10.1.1.1 1812
[Device-radius-2000] primary accounting 10.1.1.2 1813
[Device-radius-2000] key authentication abc
[Device-radius-2000] key accounting abc
[Device-radius-2000] user-name-format without-domain
[Device-radius-2000] quit
```

Create an ISP domain and specify the RADIUS scheme 2000 as the default AAA schemes for the domain.

```
[Device] domain 2000
[Device-isp-2000] authentication default radius-scheme 2000
[Device-isp-2000] authorization default radius-scheme 2000
[Device-isp-2000] accounting default radius-scheme 2000
[Device-isp-2000] quit
```

Configure a time range **ftp** for the weekdays from 8:00 to 18:00.

```
[Device] time-range ftp 8:00 to 18:00 working-day
```

Configure ACL 3000 to deny packets destined for the FTP server at 10.0.0.1 on the weekdays during business hours.

```
[Device] acl number 3000
[Device-acl-adv-3000] rule 0 deny ip destination 10.0.0.1 0 time-range ftp
[Device-acl-adv-3000] quit
```

Enable 802.1X globally.

```
[Device] dot1x
```

Enable 802.1X on port GigabitEthernet 1/0/1.

```
[Device] interface gigabitethernet 1/0/1
[Device-GigabitEthernet1/0/1] dot1x
```

Verifying the configuration

Use the user account to pass authentication, and then ping the FTP server on any weekday during business hours.

```
C:\>ping 10.0.0.1
```

```
Pinging 10.0.0.1 with 32 bytes of data:
```

```
Request timed out.
Request timed out.
Request timed out.
Request timed out.
```

```
Ping statistics for 10.0.0.1:
```

```
    Packets: Sent = 4, Received = 0, Lost = 4 (100% loss),
```

The output shows that ACL 3000 has taken effect on the user, and the user cannot access the FTP server.

Configuring EAD fast deployment

Overview

Endpoint Admission Defense (EAD) is an HP integrated endpoint access control solution, which enables the security client, security policy server, access device, and third-party server to work together to improve the threat defensive capability of a network. If a terminal device seeks to access an EAD network, it must have an EAD client, which performs 802.1X authentication.

EAD fast deployment enables the access device to redirect a user seeking to access the network to download and install EAD client. This function eliminates the tedious job of the administrator to deploy EAD clients.

EAD fast deployment is implemented by the following functions:

- [Free IP](#)
- [URL redirection](#)

Free IP

A free IP is a freely accessible network segment, which has a limited set of network resources such as software and DHCP servers. An unauthenticated user can access only this segment to download EAD client, obtain a dynamic IP address from a DHCP server, or perform some other tasks to be compliant with the network security strategy.

URL redirection

If an unauthenticated 802.1X user is using a web browser to access the network, the EAD fast deployment function redirects the user to a specific URL, for example, the EAD client software download page.

The server that provides the URL must be on the free IP accessible to unauthenticated users.

Configuration prerequisites

- Enable 802.1X globally.
- Enable 802.1X on the port, and set the port authorization mode to **auto**.

Configuring a free IP

Follow these guidelines when you configure a free IP:

- When a free IP is configured, the EAD fast deployment is enabled. To allow a user to obtain a dynamic IP address before passing 802.1X authentication, make sure the DHCP server is on the free IP segment.
- When global MAC authentication, Layer-2 portal authentication, or port security is enabled, the free IP does not take effect.
- If you use free IP, guest VLAN, and Auth-Fail VLAN features together, make sure that the free IP segments are in both guest VLAN and Auth-Fail VLAN. Users can access only the free IP segments.

To configure a free IP:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure a free IP.	dot1x free-ip <i>ip-address</i> { <i>mask-address</i> <i>mask-length</i> }	By default, no free IP is configured.

Configuring the redirect URL

Follow these guidelines when you configure the redirect URL:

- The redirect URL must be on the free IP subnet.

To configure a redirect URL:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the redirect URL.	dot1x url <i>url-string</i>	By default, no redirect URL is configured.

Setting the EAD rule timer

EAD fast deployment automatically creates an ACL rule, or an EAD rule, to open access to the redirect URL for each redirected user seeking to access the network. The EAD rule timer sets the lifetime of each ACL rule. When the timer expires or the user passes authentication, the rule is removed. If users fail to download EAD client or fail to pass authentication before the timer expires, they must reconnect to the network to access the free IP.

To prevent ACL rule resources from being used up, you can shorten the timer when the amount of EAD users is large.

To set the EAD rule timer:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the EAD rule timer.	dot1x timer ead-timeout <i>ead-timeout-value</i>	Optional. The default timer is 30 minutes.

Displaying and maintaining EAD fast deployment

Task	Command	Remarks
Display 802.1X session information, statistics, or configuration information.	display dot1x [sessions statistics] [interface <i>interface-list</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view

EAD fast deployment configuration example

Network requirements

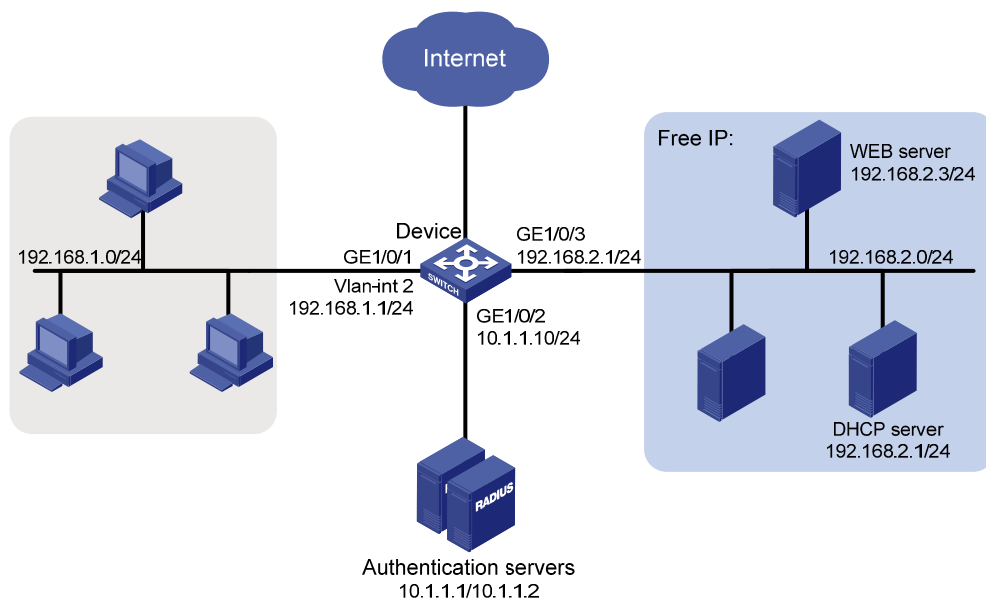
As shown in Figure 47, the hosts on the intranet 192.168.1.0/24 are attached to port GigabitEthernet 1/0/1 of the network access device, and they use DHCP to obtain IP addresses.

Deploy EAD solution for the intranet so that all hosts must pass 802.1X authentication to access the network.

To allow all intranet users to install and update 802.1X client program from a web server, configure the following:

- Allow unauthenticated users to access the segment of 192.168.2.0/24, and to obtain IP address on the segment of 192.168.1.0/24 through DHCP.
- Redirect unauthenticated users to a preconfigured web page when the users use a web browser to access any external network except 192.168.2.0/24. The web page allows users to download the 802.1X client program.
- Allow authenticated 802.1X users to access the network.

Figure 47 Network diagram



In addition to the configuration on the access device, complete the following tasks:

- Configure the DHCP server so that the host can obtain an IP address on the segment of 192.168.1.0/24.
- Configure the web server so that users can log in to the web page to download 802.1X clients.

- Configure the authentication server to provide authentication, authorization, and accounting services.

Configuration procedure

1. Configure an IP address for each interface. (Details not shown.)
2. Configure DHCP relay:

Enable DHCP.

```
<Device> system-view
```

```
[Device] dhcp enable
```

Configure a DHCP server for a DHCP server group.

```
[Device] dhcp relay server-group 1 ip 192.168.2.2
```

Enable the relay agent on VLAN interface 2.

```
[Device] interface vlan-interface 2
```

```
[Device-Vlan-interface2] dhcp select relay
```

Correlate VLAN interface 2 to the DHCP server group.

```
[Device-Vlan-interface2] dhcp relay server-select 1
```

```
[Device-Vlan-interface2] quit
```

3. Configure a RADIUS scheme and an ISP domain.

For more information about configuration procedure, see "Configuring 802.1X."

4. Configure 802.1X:

Configure the free IP.

```
[Device] dot1x free-ip 192.168.2.0 24
```

Configure the redirect URL for client software download.

```
[Device] dot1x url http://192.168.2.3
```

Enable 802.1X globally.

```
[Device] dot1x
```

Enable 802.1X on the port.

```
[Device] interface gigabitethernet 1/0/1
```

```
[Device-GigabitEthernet1/0/1] dot1x
```

Verifying the configuration

Use the **display dot1x** command to display the 802.1X configuration. After the host obtains an IP address from a DHCP server, use the **ping** command from the host to ping an IP address on the network segment specified by free IP.

```
C:\>ping 192.168.2.3
```

```
Pinging 192.168.2.3 with 32 bytes of data:
```

```
Reply from 192.168.2.3: bytes=32 time<1ms TTL=128
```

```
Reply from 192.168.2.3: bytes=32 time<1ms TTL=128
```

```
Reply from 192.168.2.3: bytes=32 time<1ms TTL=128
```

```
Reply from 192.168.2.3: bytes=32 time<1ms TTL=128
```

```
Ping statistics for 192.168.2.3:
```

```
Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),  
Approximate round trip times in milli-seconds:  
Minimum = 0ms, Maximum = 0ms, Average = 0ms
```

The output shows that you can access that segment before passing 802.1X authentication. If you use a web browser to access any external website beyond the free IP segments, you are redirected to the web server, which provides the 802.1X client software download service. Enter the external website address in dotted decimal notation, for example, 3.3.3.3 or <http://3.3.3.3>, in the address bar.

Troubleshooting EAD fast deployment

Web browser users cannot be correctly redirected

Symptom

Unauthenticated users are not redirected to the specified redirect URL after they enter external website addresses in their web browsers.

Analysis

Redirection will not happen for one of the following reasons:

- The address is in the string format. The operating system of the host regards the string as a website name and tries to resolve it. If the resolution fails, the operating system sends an ARP request, but the target address is not in the dotted decimal notation. The redirection function does not redirect this kind of ARP request.
- The address is within a free IP segment. No redirection will take place, even if no host is present with the address.
- The redirect URL is not in a free IP segment, no server is using the redirect URL, or the server with the URL does not provide web services.

Solution

1. Enter a dotted decimal IP address that is not in any free IP segment.
2. Make sure that the network access device and the server are correctly configured.

Configuring MAC authentication

MAC authentication overview

MAC authentication controls network access by authenticating source MAC addresses on a port. It does not require client software. A user does not need to input a username and password for network access. The device initiates a MAC authentication process when it detects an unknown source MAC address on a MAC authentication enabled port. If the MAC address passes authentication, the user can access authorized network resources. If the authentication fails, the device marks the MAC address as a silent MAC address, drops the packet, and starts a quiet timer. The device drops all subsequent packets from the MAC address within the quiet time. This quiet mechanism avoids repeated authentication during a short time.

NOTE:

If the MAC address that has failed authentication is a static MAC address or a MAC address that has passed any security authentication, the device does not mark it as a silent address.

User account policies

MAC authentication supports the following user account policies:

- One MAC-based user account for each user. The access device uses the source MAC addresses in packets as the usernames and passwords of users for MAC authentication. This policy is suitable for an insecure environment.
- One shared user account for all users. You specify one username and password, which are not necessarily a MAC address, for all MAC authentication users on the access device. This policy is suitable for a secure environment.

Authentication approaches

You can perform MAC authentication on the access device (local authentication) or through a Remote Authentication Dial-In User Service (RADIUS) server.

Suppose a source MAC unknown packet arrives at a MAC authentication enabled port.

In the local authentication approach:

- If MAC-based accounts are used, the access device uses the source MAC address of the packet as the username and password to search its local account database for a match.
- If a shared account is used, the access device uses the shared account username and password to search its local account database for a match.

In the RADIUS authentication approach:

- If MAC-based accounts are used, the access device sends the source MAC address as the username and password to the RADIUS server for authentication.
- If a shared account is used, the access device sends the shared account username and password to the RADIUS server for authentication.

For more information about configuring local authentication and RADIUS authentication, see "[Configuring AAA](#)."

MAC authentication timers

MAC authentication uses the following timers:

- **Offline detect timer**—Sets the interval that the device waits for traffic from a user before it regards the user idle. If a user connection has been idle for two consecutive intervals, the device logs the user out and stops accounting for the user.
- **Quiet timer**—Sets the interval that the device must wait before it can perform MAC authentication for a user that has failed MAC authentication. All packets from the MAC address are dropped during the quiet time. This quiet mechanism prevents repeated authentication from affecting system performance.
- **Server timeout timer**—Sets the interval that the access device waits for a response from a RADIUS server before it regards the RADIUS server unavailable. If the timer expires during MAC authentication, the user cannot access the network.

Using MAC authentication with other features

VLAN assignment

You can specify a VLAN in the user account for a MAC authentication user to control the account's access to network resources. After the user passes MAC authentication, the authentication server, either the local access device or a RADIUS server, assigns the VLAN to the port as the default VLAN. After the user logs off, the initial default VLAN, or the default VLAN configured before any VLAN is assigned by the authentication server, restores. If the authentication server assigns no VLAN, the initial default VLAN applies.

A hybrid port is always assigned to a server-assigned VLAN as an untagged member. After the assignment, do not re-configure the port as a tagged member in the VLAN.

If MAC-based VLAN is enabled on a hybrid port, the device maps the server-assigned VLAN to the MAC address of the user. The default VLAN of the hybrid port does not change.

ACL assignment

You can specify an ACL in the user account for a MAC authentication user to control its access to network resources. After the user passes MAC authentication, the authentication server, either the local access device or a RADIUS server, assigns the ACL to the access port to filter the traffic from this user. You must configure the ACL on the access device for the ACL assignment function. You can change ACL rules while the user is online.

Guest VLAN

You can configure a guest VLAN to accommodate MAC authentication users that have failed MAC authentication on the port. Users in the MAC authentication guest VLAN can access a limited set of network resources, such as a software server, to download anti-virus software and system patches. If no MAC authentication guest VLAN is configured, the user that fails MAC authentication cannot access any network resources.

If a user in the guest VLAN passes MAC authentication, it is removed from the guest VLAN and can access all authorized network resources. If not, the user is still in the MAC authentication guest VLAN.

A hybrid port is always assigned to a guest VLAN as an untagged member. After the assignment, do not re-configure the port as a tagged member in the VLAN.

Critical VLAN

You can configure a MAC authentication critical VLAN on a port to accommodate users that fail MAC authentication because no RADIUS authentication server is reachable. Users in a MAC authentication critical VLAN can access a limit set of network resources depending on your configuration.

The critical VLAN feature takes effect when MAC authentication is performed only through RADIUS servers. If a MAC authentication user fails local authentication after RADIUS authentication, the user is not assigned to the critical VLAN. For more information about RADIUS configuration, see "[Configuring AAA](#)."

Any of the following RADIUS authentication server changes in the ISP domain for MAC authentication users on a port can cause users to be removed from the critical VLAN:

- An authentication server is reconfigured, added, or removed.
- The status of any RADIUS authentication server automatically changes to active or is administratively set to active.
- The RADIUS server probing function detects that a RADIUS authentication server is reachable and sets its state to active.

Configuration task list

Task	Remarks	
Basic configuration for MAC authentication	Configuring MAC authentication globally	Required
	Configuring MAC authentication on a port	Required
Specifying a MAC authentication domain	Optional	
Configuring a MAC authentication guest VLAN	Optional	
Configuring a MAC authentication critical VLAN	Optional	

Basic configuration for MAC authentication

- Create and configure an authentication domain, also called "an ISP domain."
- For local authentication, create local user accounts, and specify the **lan-access** service for the accounts.
- For RADIUS authentication, check that the device and the RADIUS server can reach each other, and create user accounts on the RADIUS server.

If you are using MAC-based accounts, make sure that the username and password for each account is the same as the MAC address of the MAC authentication users.

MAC authentication can take effect on a port only when it is enabled globally and on the port.

Configuring MAC authentication globally

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable MAC authentication globally.	mac-authentication	Disabled by default.
3. Configure MAC authentication timers.	mac-authentication timer { offline-detect <i>offline-detect-value</i> quiet <i>quiet-value</i> server-timeout <i>server-timeout-value</i> }	Optional. By default, the offline detect timer is 300 seconds, the quiet timer is 60 seconds, and the server timeout timer is 100 seconds.
4. Configure the properties of MAC authentication user accounts.	mac-authentication user-name-format { fixed [account <i>name</i>] [password { cipher simple } <i>password</i>] mac-address [{ with-hyphen without-hyphen } [lowercase uppercase]] }	Optional. By default, the username and password for a MAC authentication user account must be a MAC address in lower case without hyphens.

NOTE:

When global MAC authentication is enabled, the EAD fast deployment function cannot take effect.

Configuring MAC authentication on a port

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable MAC authentication.	<ul style="list-style-type: none"> • (Approach 1) In system view: mac-authentication interface <i>interface-list</i> • (Approach 2) In interface view: <ol style="list-style-type: none"> a. interface <i>interface-type</i> <i>interface-number</i> b. mac-authentication 	Disabled by default. Enable MAC authentication for ports in bulk in system view or an individual port in Ethernet interface view.
3. Set the maximum number of concurrent MAC authentication users allowed on a port.	mac-authentication max-user <i>user-number</i>	Optional. By default, the maximum number of concurrent MAC authentication users is 256.

NOTE:

You cannot add a MAC authentication enabled port in to a link aggregation group, or enable MAC authentication on a port already in a link aggregation group.

Specifying a MAC authentication domain

By default, MAC authentication users are in the system default authentication domain. To implement different access policies for users, you can specify authentication domains for MAC authentication users in the following ways:

- Specify a global authentication domain in system view. This domain setting applies to all ports.
- Specify an authentication domain for an individual port in Ethernet interface view.

MAC authentication chooses an authentication domain for users on a port in this order: the interface-specific domain, the global domain, and the default domain. For more information about authentication domains, see "[Configuring AAA](#)."

To specify an authentication domain for MAC authentication users:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Specify an authentication domain for MAC authentication users.	<ul style="list-style-type: none">• (Approach 1) In system view: mac-authentication domain <i>domain-name</i>• (Approach 2) In interface view:<ul style="list-style-type: none">a. interface <i>interface-type</i> <i>interface-number</i>b. mac-authentication domain <i>domain-name</i>	Use either approach. By default, the system default authentication domain is used for MAC authentication users.

Configuring a MAC authentication guest VLAN

Before you configure a MAC authentication guest VLAN on a port, complete the following tasks:

- Enable MAC authentication.
- Enable MAC-based VLAN on the port.
- Create the VLAN to be specified as the MAC authentication guest VLAN.

To configure a MAC authentication guest VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Ethernet port view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Specify a MAC authentication guest VLAN.	mac-authentication guest-vlan <i>guest-vlan-id</i>	By default, no MAC authentication guest VLAN is configured. You can configure only one MAC authentication guest VLAN on a port.

Follow the guidelines in [Table 8](#) when configuring a MAC authentication guest VLAN on a port.

Table 8 Relationships of the MAC authentication guest VLAN with other security features

Feature	Relationship description	Reference
Quiet function of MAC authentication	The MAC authentication guest VLAN function has higher priority. A user can access any resources in the guest VLAN.	See " MAC authentication timers "
Super VLAN	You cannot specify a VLAN as both a super VLAN and a MAC authentication guest VLAN.	See <i>Layer 2—LAN Switching Configuration Guide</i>
Port intrusion protection	The MAC authentication guest VLAN function has higher priority than the block MAC action but lower priority than the shut down port action of the port intrusion protection feature.	See " Configuring port security "
802.1X guest VLAN on a port that performs MAC-based access control	The MAC authentication guest VLAN has a lower priority.	See " Configuring 802.1X "

Configuring a MAC authentication critical VLAN

Before you configure a MAC authentication critical VLAN on a port, complete the following tasks:

- Enable MAC authentication.
- Enable MAC-based VLAN on the port.
- Create the VLAN to be specified as the MAC authentication critical VLAN.

To configure a MAC authentication critical VLAN:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet port view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Specify a MAC authentication critical VLAN.	mac-authentication critical vlan <i>critical-vlan-id</i>	By default, no MAC authentication critical VLAN is configured. You can configure only one MAC authentication critical VLAN on a port.

Follow the guidelines in [Table 9](#) when you configure a MAC authentication critical VLAN on a port.

Table 9 Relationships of the MAC authentication critical VLAN with other security features

Feature	Relationship description	Reference
Quiet function of MAC authentication	The MAC authentication critical VLAN function has higher priority. When a user fails MAC authentication because no RADIUS authentication server is reachable, the user can access the resources in the critical VLAN, and the user's MAC address is not marked as a silent MAC address.	See " MAC authentication timers "
Super VLAN	You cannot specify a VLAN as both a super VLAN and a MAC authentication critical VLAN.	See <i>Layer 2—LAN Switching Configuration Guide</i>
Port intrusion protection	The MAC authentication critical VLAN function has higher priority than the block MAC action but lower priority than the shut down port action of the port intrusion protection feature.	See " Configuring port security "

Displaying and maintaining MAC authentication

Task	Command	Remarks
Display MAC authentication information.	display mac-authentication [interface <i>interface-list</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear MAC authentication statistics.	reset mac-authentication statistics [interface <i>interface-list</i>]	Available in user view

MAC authentication configuration examples

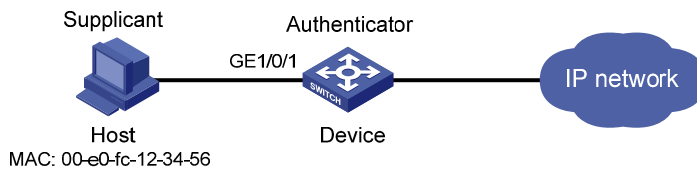
Local MAC authentication configuration example

Network requirements

In the network in [Figure 48](#), perform local MAC authentication on port GigabitEthernet 1/0/1 to control Internet access. Make sure that:

- All users belong to domain aabbcc.net.
- Local users use their MAC address as the username and password for MAC authentication. The MAC addresses are hyphen separated and in lower case.
- The access device detects whether a user has gone offline every 180 seconds. When a user fails authentication, the device does not authenticate the user within 180 seconds.

Figure 48 Network diagram



Configuration procedure

Add a local user account, set both the username and password to 00-e0-fc-12-34-56, the MAC address of the user host, and enable LAN access service for the account.

```
<Device> system-view
[Device] local-user 00-e0-fc-12-34-56
[Device-luser-00-e0-fc-12-34-56] password simple 00-e0-fc-12-34-56
[Device-luser-00-e0-fc-12-34-56] service-type lan-access
[Device-luser-00-e0-fc-12-34-56] quit
```

Configure ISP domain **aabbcc.net** to perform local authentication for LAN access users.

```
[Device] domain aabbcc.net
[Device-isp-aabbcc.net] authentication lan-access local
[Device-isp-aabbcc.net] quit
```

Enable MAC authentication globally.

```
[Device] mac-authentication
```

Enable MAC authentication on port GigabitEthernet 1/0/1.

```
[Device] mac-authentication interface gigabitethernet 1/0/1
```

Specify the ISP domain for MAC authentication.

```
[Device] mac-authentication domain aabbcc.net
```

Set the MAC authentication timers.

```
[Device] mac-authentication timer offline-detect 180
[Device] mac-authentication timer quiet 180
```

Configure MAC authentication to use MAC-based accounts. The MAC address usernames and passwords are hyphenated and in lowercase.

```
[Device] mac-authentication user-name-format mac-address with-hyphen lowercase
```

Verifying the configuration

Display MAC authentication settings and statistics.

```
<Device> display mac-authentication
MAC address authentication is enabled.
User name format is MAC address in lowercase, like xx-xx-xx-xx-xx-xx
Fixed username:mac
Fixed password:not configured
    Offline detect period is 180s
    Quiet period is 180s.
    Server response timeout value is 100s
    The max allowed user number is 1024 per slot
    Current user number amounts to 1
    Current domain is aabbcc.net
Silent Mac User info:
```

```

MAC Addr          From Port          Port Index
GigabitEthernet1/0/1 is link-up
MAC address authentication is enabled
Authenticate success: 1, failed: 0
Max number of on-line users is 256
Current online user number is 1
MAC Addr          Authenticate state      Auth Index
00e0-fc12-3456    MAC_AUTHENTICATOR_SUCCESS    29

```

After the user passes authentication, use the **display connection** command to display the online user information.

```

<Device> display connection
Slot: 1
Index=29 ,Username=00-e0-fc-12-34-56@aabbcc.net
IP=N/A
IPv6=N/A
MAC=00e0-fc12-3456
Total 1 connection(s) matched on slot 1.
Total 1 connection(s) matched.

```

RADIUS-based MAC authentication configuration example

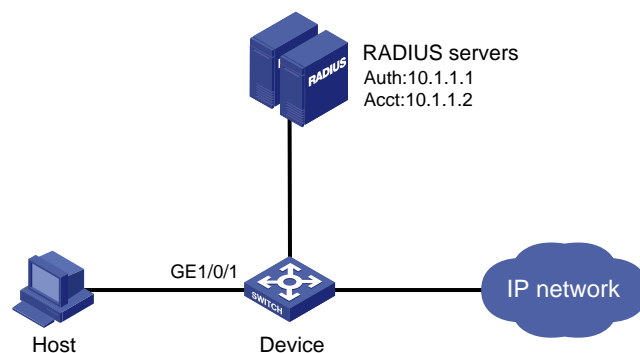
Network requirements

As shown in [Figure 49](#), a host connects to port GigabitEthernet 1/0/1 on the access device. The device uses RADIUS servers for authentication, authorization, and accounting.

Perform MAC authentication on port GigabitEthernet 1/0/1 to control Internet access. Make sure that:

- The device detects whether a user has gone offline every 180 seconds. If a user fails authentication, the device does not authenticate the user within 180 seconds.
- All MAC authentication users belong to ISP domain 2000 and share the user account **aaa** with password **123456**.

Figure 49 Network diagram



Configuration procedure

1. Make sure the RADIUS server and the access device can reach each other.
2. Create a shared account for MAC authentication users on the RADIUS server, and set the username **aaa** and password **123456** for the account.
3. Configure the device:

```

# Configure a RADIUS scheme.
<Device> system-view
[Device] radius scheme 2000
[Device-radius-2000] primary authentication 10.1.1.1 1812
[Device-radius-2000] primary accounting 10.1.1.2 1813
[Device-radius-2000] key authentication abc
[Device-radius-2000] key accounting abc
[Device-radius-2000] user-name-format without-domain
[Device-radius-2000] quit

# Apply the RADIUS scheme to ISP domain 2000 for authentication, authorization, and
accounting.
[Device] domain 2000
[Device-isp-2000] authentication default radius-scheme 2000
[Device-isp-2000] authorization default radius-scheme 2000
[Device-isp-2000] accounting default radius-scheme 2000
[Device-isp-2000] quit

# Enable MAC authentication globally.
[Device] mac-authentication

# Enable MAC authentication on port GigabitEthernet 1/0/1.
[Device] mac-authentication interface gigabitethernet 1/0/1

# Specify the ISP domain for MAC authentication.
[Device] mac-authentication domain 2000

# Set the MAC authentication timers.
[Device] mac-authentication timer offline-detect 180
[Device] mac-authentication timer quiet 180

# Specify username aaa and plaintext password 123456 for the account shared by MAC
authentication users.
[Device] mac-authentication user-name-format fixed account aaa password simple 123456

```

Verifying the configuration

```

# Display MAC authentication settings and statistics.
<Device> display mac-authentication
MAC address authentication is enabled.
User name format is fixed account
Fixed username:aaa
Fixed password: *****
    Offline detect period is 180s
    Quiet period is 180s.
    Server response timeout value is 100s
    The max allowed user number is 1024 per slot
    Current user number amounts to 1
    Current domain is 2000
Silent Mac User info:
    MAC ADDR           From Port           Port Index
Gigabitethernet1/0/1 is link-up
    MAC address authentication is enabled
    Authenticate success: 1, failed: 0

```

```

Max number of on-line users is 256
Current online user number is 1
MAC ADDR           Authenticate state           Auth Index
00e0-fc12-3456    MAC_AUTHENTICATOR_SUCCESS         29

```

After a user passes MAC authentication, use the **display connection** command to display online user information.

```

<Device> display connection
Slot: 1
Index=29 ,Username=aaa@2000
IP=N/A
IPv6=N/A
MAC=00e0-fc12-3456
Total 1 connection(s) matched on slot 1.
Total 1 connection(s) matched.

```

ACL assignment configuration example

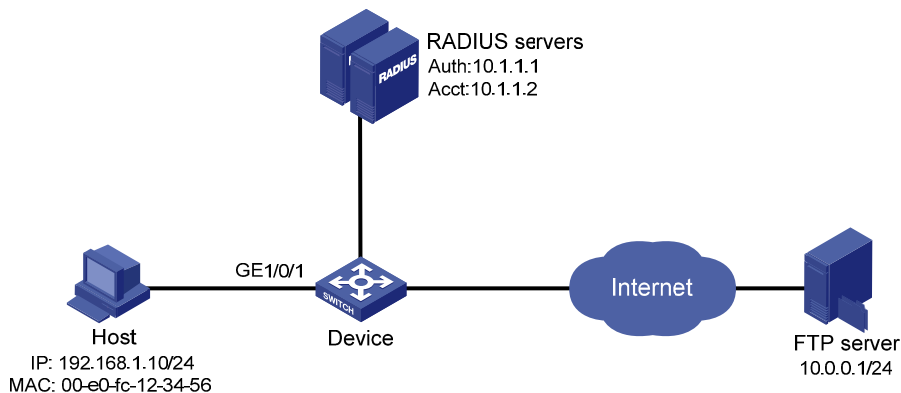
Network requirements

As shown in [Figure 50](#), a host connects to the device's port GigabitEthernet 1/0/1, and the device uses RADIUS servers to perform authentication, authorization, and accounting.

Perform MAC authentication on port GigabitEthernet 1/0/1 to control Internet access. Make sure that an authenticated user can access the Internet but the FTP server at 10.0.0.1.

Use MAC-based user accounts for MAC authentication users. The MAC addresses are hyphen separated and in lower case.

Figure 50 Network diagram



Configuration procedure

1. Make sure the RADIUS server and the access device can reach each other.
2. Configure the ACL assignment:

```
# Configure ACL 3000 to deny packets destined for 10.0.0.1.
```

```

<Sysname> system-view
[Sysname] acl number 3000
[Sysname-acl-adv-3000] rule 0 deny ip destination 10.0.0.1 0
[Sysname-acl-adv-3000] quit

```

3. Configure RADIUS-based MAC authentication on the device:

Configure a RADIUS scheme.

```
[Sysname] radius scheme 2000
[Sysname-radius-2000] primary authentication 10.1.1.1 1812
[Sysname-radius-2000] primary accounting 10.1.1.2 1813
[Sysname-radius-2000] key authentication simple abc
[Sysname-radius-2000] key accounting simple abc
[Sysname-radius-2000] user-name-format without-domain
[Sysname-radius-2000] quit
```

Apply the RADIUS scheme to an ISP domain for authentication, authorization, and accounting.

```
[Sysname] domain 2000
[Sysname-isp-2000] authentication default radius-scheme 2000
[Sysname-isp-2000] authorization default radius-scheme 2000
[Sysname-isp-2000] accounting default radius-scheme 2000
[Sysname-isp-2000] quit
```

Enable MAC authentication globally.

```
[Sysname] mac-authentication
```

Specify the ISP domain for MAC authentication.

```
[Sysname] mac-authentication domain 2000
```

Configure the device to use MAC-based user accounts, and the MAC addresses are hyphen separated and in lowercase.

```
[Sysname] mac-authentication user-name-format mac-address with-hyphen lowercase
```

Enable MAC authentication for port GigabitEthernet 1/0/1.

```
[Sysname] interface gigabitethernet 1/0/1
[Sysname-GigabitEthernet1/0/1] mac-authentication
```

4. Configure the RADIUS servers:

Add a user account with **00-e0-fc-12-34-56** as both the username and password on the RADIUS server, and specify ACL 3000 as the authorization ACL for the user account. (Details not shown.)

Verifying the configuration

After the host passes authentication, perform the **display connection** command on the device to view online user information.

```
[Sysname-GigabitEthernet1/0/1] display connection
Slot: 1
Index=9 , Username=00-e0-fc-12-34-56@2000
IP=N/A
IPv6=N/A
MAC=00e0-fc12-3456
```

```
Total 1 connection(s) matched on slot 1.
```

```
Total 1 connection(s) matched.
```

Ping the FTP server from the host to verify that the ACL 3000 has been assigned to port GigabitEthernet 1/0/1 to deny access to the FTP server.

```
C:\>ping 10.0.0.1
```

```
Pinging 10.0.0.1 with 32 bytes of data:
```


Request timed out.

Request timed out.

Request timed out.

Request timed out.

Ping statistics for 10.0.0.1:

Packets: Sent = 4, Received = 0, Lost = 4 (100% loss),

Configuring portal authentication

The IPv6 portal configuration is available only on the HP 5500 EI switch series.

Overview

Portal authentication helps control access to the Internet. It is also called "web authentication." A website implementing portal authentication is called a portal website.

With portal authentication, an access device redirects all users to the portal authentication page. All users can access the free services provided on the portal website; but to access the Internet, a user must pass portal authentication.

A user can access a known portal website and enter a username and password for authentication. This authentication mode is called active authentication. There is another authentication mode, forced authentication, in which the access device forces a user who is trying to access the Internet through Hypertext Transfer Protocol (HTTP) to log on to a portal website for authentication.

The portal feature provides the flexibility for Internet service providers (ISPs) to manage services. A portal website can, for example, present advertisements and deliver community and personalized services. In this way, broadband network providers, equipment vendors, and content service providers form an industrial ecological system.

Extended portal functions

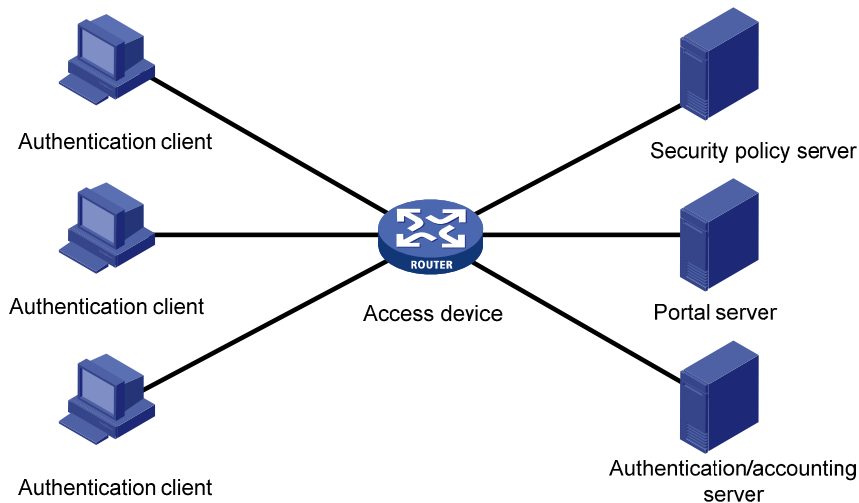
By forcing patching and anti-virus policies, extended portal functions help users to defend against viruses. Portal authentication supports the following extended functions:

- **Security check**—Works after identity authentication succeeds to check whether the required anti-virus software, virus definition file, and operating system patches are installed, and whether there is any unauthorized software installed on the user host.
- **Resource access restriction**—Allows users passing identity authentication to access only network resources in the quarantined area, such as the anti-virus server and the patch server. Only users passing both identity authentication and security check can access restricted network resources.

Portal system components

A typical portal system comprises these basic components: authentication client, access device, portal server, authentication/accounting server, and security policy server.

Figure 51 Portal system components



Authentication client

An authentication client is an entity seeking access to network resources. It is typically an end-user terminal, such as a PC. A client can use a browser or a portal client software for portal authentication. Client security check is implemented through communications between the client and the security policy server.

Access device

Access devices control user access. An access device can be a switch or router that provides the following functions:

- Redirecting all HTTP requests from unauthenticated users to the portal server.
- Interacting with the portal server, the security policy server, and the authentication/accounting server for identity authentication, security check, and accounting.
- Allowing users who have passed identity authentication and security check to access granted Internet resources.

Portal server

A portal server listens to authentication requests from authentication clients and exchanges client authentication information with the access device. It provides free portal services and pushes web authentication pages to users.

NOTE:

A portal server can be an entity independent of the access device or an entity embedded in the access device. In this document, the term portal server refers to an independent portal server, and the term local portal server refers to an embedded portal server. Only the HP 5500 EI series support an independent portal server.

Authentication/accounting server

An authentication/accounting server implements user authentication and accounting through interaction with the access device.

Only a RADIUS server can serve as the remote authentication/accounting server in a portal system.

Security policy server

A security policy server interacts with authentication clients and access devices for security check and resource authorization.

The components of a portal system interact in the following procedure:

1. When an unauthenticated user enters a website address in the browser's address bar to access the Internet, an HTTP request is created and sent to the access device, which redirects the HTTP request to the portal server's web authentication homepage. For extended portal functions, authentication clients must run the portal client software.
2. On the authentication homepage/authentication dialog box, the user enters and submits the authentication information, which the portal server then transfers to the access device.
3. Upon receipt of the authentication information, the access device communicates with the authentication/accounting server for authentication and accounting.
4. After successful authentication, the access device checks whether there is a corresponding security policy for the user. If not, it allows the user to access the Internet. Otherwise, the client communicates with the access device and the security policy server for security check. If the client passes security check, the security policy server authorizes the user to access the Internet resources.

NOTE:

To implement security check, the client must be the HP iNode client.

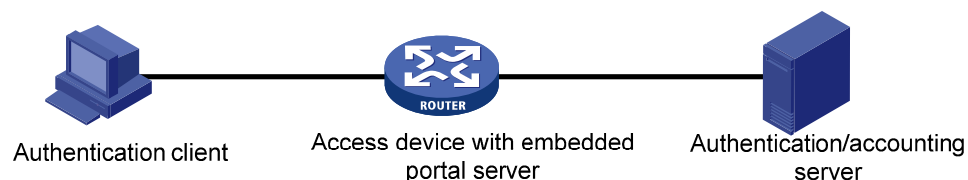
Portal authentication supports NAT traversal whether it is initiated by a web client or an HP iNode client. When the portal authentication client is on a private network, but the portal server is on a public network and the access device is enabled with NAT, network address translations performed on the access device do not affect portal authentication. However, in such a case, HP recommends using an interface's public IP address as the source address of outgoing portal packets.

Portal system using the local portal server

System components

In addition to use a separate device as the portal server, a portal system can also use the local portal server function of the access device to authenticate web users directly. A portal system using the local portal server does not support extended portal functions. No security policy server is needed for local portal service. In this case, the portal system consists of only three components: authentication client, access device, and authentication/accounting server, as shown in [Figure 52](#).

Figure 52 Portal system using the local portal server



NOTE:

The local portal server function of the access device implements only some simple portal server functions. It only allows users to log on and log off through the web interface. It cannot take the place of an independent portal server.

Protocols used for interaction between the client and local portal server

HTTP and Hypertext Transfer Protocol Secure (HTTPS) can be used for interaction between an authentication client and an access device providing the local portal server function. If HTTP is used, there are potential security problems because HTTP packets are transferred in plain text; if HTTPS is used, secure data transmission is ensured because HTTPS packets are transferred in cipher text based on SSL.

Authentication page customization support

The local portal server function allows you to customize authentication pages. You can customize authentication pages by editing the corresponding HTML files and then compress and save the files to the storage medium of the device. A set of customized authentication pages consists of six authentication pages—the logon page, the logon success page, the online page, the logoff success page, the logon failure page, and the system busy page. A local portal server will push a corresponding authentication page at each authentication phase. If you do not customize the authentication pages, the local portal server will push the default authentication pages.

For the rules of customizing authentication pages, see "[Customizing authentication pages.](#)"

Portal authentication modes

Portal authentication may work at Layer 2 or Layer 3 of the OSI model.

Layer 2 portal authentication

You can enable Layer 2 portal authentication on an access device's Layer 2 ports that connect authentication clients, so that only clients whose MAC addresses pass authentication can access the external network. Only the local portal server provided by the access device supports Layer 2 portal authentication.

Layer 2 portal authentication allows the authentication server to assign different VLANs according to user authentication results so that access devices can thereby control user access to resources. After a client passes authentication, the authentication server can assign an authorized VLAN to allow the user to access the resources in the VLAN. If a client fails authentication, the authentication server can assign an Auth-Fail VLAN. Layer 3 portal authentication does not support VLAN assignment.

Layer 3 portal authentication (available only on the HP 5500 EI series)

You can enable Layer 3 authentication on an access device's Layer 3 interfaces that connect authentication clients. Portal authentication performed on a Layer 3 interface can be direct authentication, re-DHCP authentication, or cross-subnet authentication. In direct authentication and re-DHCP authentication, no Layer-3 forwarding devices exist between the authentication client and the access device. In cross-subnet authentication, Layer-3 forwarding devices may exist between the authentication client and the access device.

- Direct authentication
Before authentication, a user manually configures a public IP address or directly obtains a public IP address through DHCP, and can access only the portal server and predefined free websites. After passing authentication, the user can access the network resources. The process of direct authentication is simpler than that of re-DHCP authentication.
- Re-DHCP authentication
Before authentication, a user gets a private IP address through DHCP and can access only the portal server and predefined free websites. After passing authentication, the user is allocated a public IP address and can access the network resources. No public IP address is allocated to those who fail authentication. This solves the IP address planning and allocation problem and can be

useful. For example, a service provider can allocate public IP addresses to broadband users only when they access networks beyond the residential community network.

The local portal server does not support re-DHCP portal authentication.

IPv6 portal authentication does not support the re-DHCP authentication mode.

- Cross-subnet authentication

Cross-subnet authentication is similar to direct authentication, but it allows Layer 3 forwarding devices to be present between the authentication client and the access device.

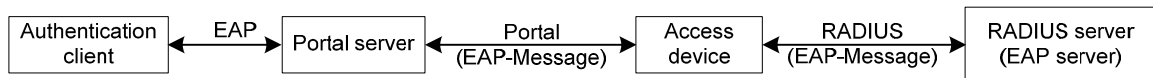
In direct authentication, re-DHCP authentication, and cross-subnet authentication, the client's IP address is used for client identification. After a client passes authentication, the access device generates an access control list (ACL) for the client based on the client's IP address to permit packets from the client to go through the access port. Because no Layer 3 devices are present between the authentication clients and the access device in direct authentication and re-DHCP authentication, the access device can directly learn the clients' MAC addresses, and can enhance the capability of controlling packet forwarding by also using the learned MAC addresses.

Portal support for EAP (available only on the HP 5500 EI series)

Authentication by using the username and password is less secure. Digital certificate authentication is usually used to ensure higher security.

The Extensible Authentication Protocol (EAP) supports several digital certificate-based authentication methods, for example, EAP-TLS. Working together with EAP, portal authentication can implement digital certificate-based user authentication.

Figure 53 Portal support for EAP working flow diagram



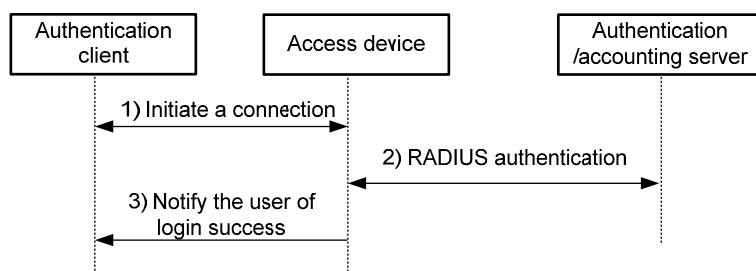
As shown in Figure 53, the authentication client and the portal server exchange EAP authentication packets. The portal server and the access device exchange portal authentication packets that carry the EAP-Message attributes. The access device and the RADIUS server exchange RADIUS packets that carry the EAP-Message attributes. The RADIUS server that supports the EAP server function processes the EAP packets encapsulated in the EAP-Message attributes, and provides the EAP authentication result. During the whole EAP authentication process, the access device does not process the packets that carry the EAP-Message attributes but only transports them between the portal server and the RADIUS server. Therefore, no additional configuration is needed on the access device.

NOTE:

- To use portal authentication that supports EAP, the portal server and client must be the IMC portal server and the iNode portal client.
 - Only Layer 3 portal authentication that uses a remote portal server supports EAP authentication.
-

Layer 2 portal authentication process

Figure 54 Local Layer 2 portal authentication process



Local Layer 2 portal authentication takes the following procedure:

1. The portal authentication client sends an HTTP or HTTPS request. Upon receiving the HTTP request, the access device redirects it to the listening IP address of the local portal server, which then pushes a web authentication page to the authentication client. The user types the username and password on the web authentication page. The listening IP address of the local portal server is the IP address of a Layer 3 interface on the access device that can communicate with the portal client. Usually, it is a loopback interface's IP address.
2. The access device and the RADIUS server exchange RADIUS packets to authenticate the user.
3. If the user passes RADIUS authentication, the local portal server pushes a logon success page to the authentication client.

Authorized VLAN

Layer 2 portal authentication supports VLAN assignment by the authentication server. After a user passes portal authentication, if the authentication server is configured with an authorized VLAN for the user, the authentication server assigns the authorized VLAN to the access device. Then, the access device adds the user to the authorized VLAN and generates a MAC VLAN entry. If the authorized VLAN does not exist, the access device first creates the VLAN.

By deploying the authorized VLAN assignment function, you can control which authenticated users can access which network resources.

Auth-Fail VLAN

The Auth-Fail VLAN feature allows users failing authentication to access a VLAN that accommodates network resources such as the patches server, virus definitions server, client software server, and anti-virus software server, so that the users can upgrade their client software or other programs. Such a VLAN is called an Auth-Fail VLAN.

Layer 2 portal authentication supports Auth-Fail VLAN on a port that performs MAC-based access control. With an Auth-Fail VLAN configured on a port, if a user on the port fails authentication, the access device creates a MAC VLAN entry based on the MAC address of the user and adds the user to the Auth-Fail VLAN. Then, the user can access the non-HTTP resources in the Auth-Fail VLAN, and all HTTP requests of the user will be redirected to the authentication page. If the user passes authentication, the access device adds the user to the assigned VLAN or return the user to the initial VLAN of the port, depending on whether the authentication server assigns a VLAN. If the user fails the authentication, the access device keeps the user in the Auth-Fail VLAN. If an access port receives no traffic from a user in the Auth-Fail VLAN during a specified period of time (90 seconds by default), it removes the user from the Auth-Fail VLAN and adds the user to the initial VLAN of the port.

NOTE:

After a user is added to the authorized VLAN or Auth-Fail VLAN, the IP address of the client needs to be automatically or manually updated to make sure that the client can communicate with the hosts in the VLAN.

Assignment of authorized ACLs

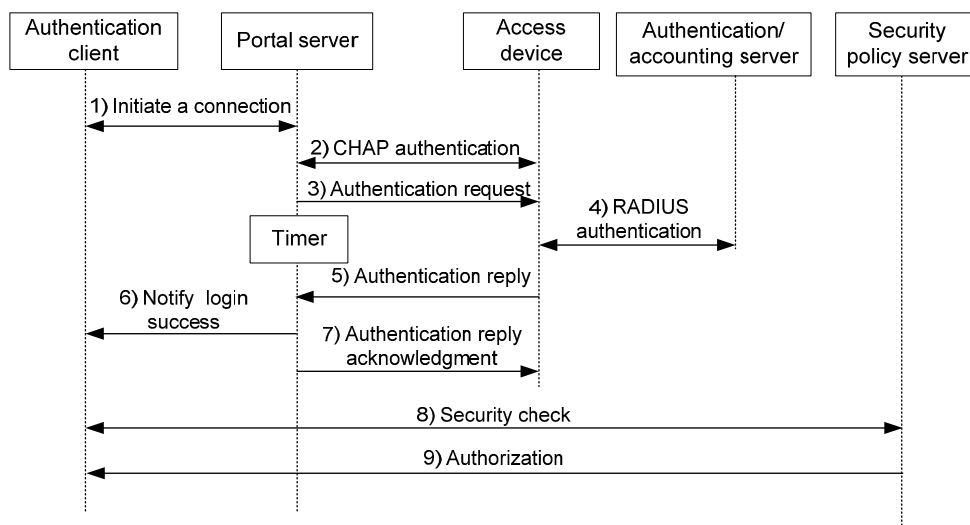
The device can use ACLs to control user access to network resources and limit user access rights. With authorized ACLs specified on the authentication server, when a user passes authentication, the authentication server assigns an authorized ACL for the user, and the device filters traffic from the user on the access port according to the authorized ACL. You must configure the authorized ACLs on the access device if you specify authorized ACLs on the authentication server. To change the access right of a user, specify a different authorized ACL on the authentication server or change the rules of the corresponding authorized ACL on the device.

Layer 3 portal authentication process (available only on the HP 5500 EI series)

Direct authentication and cross-subnet authentication share the same authentication process, while re-DHCP authentication has a different process because of the presence of two address allocation procedures.

Direct authentication/cross-subnet authentication process (with CHAP/PAP authentication)

Figure 55 Direct authentication/cross-subnet authentication process



The direct authentication/cross-subnet authentication takes the following procedure:

1. An authentication client initiates authentication by sending an HTTP request. When the HTTP packet arrives at the access device, the access device allows it to pass if it is destined for the portal server or a predefined free website, or redirects it to the portal server if it is destined for other websites. The portal server pushes a web authentication page to the user and the user enters the username and password.
2. The portal server and the access device exchange Challenge Handshake Authentication Protocol (CHAP) messages. For Password Authentication Protocol (PAP) authentication, this step is skipped.

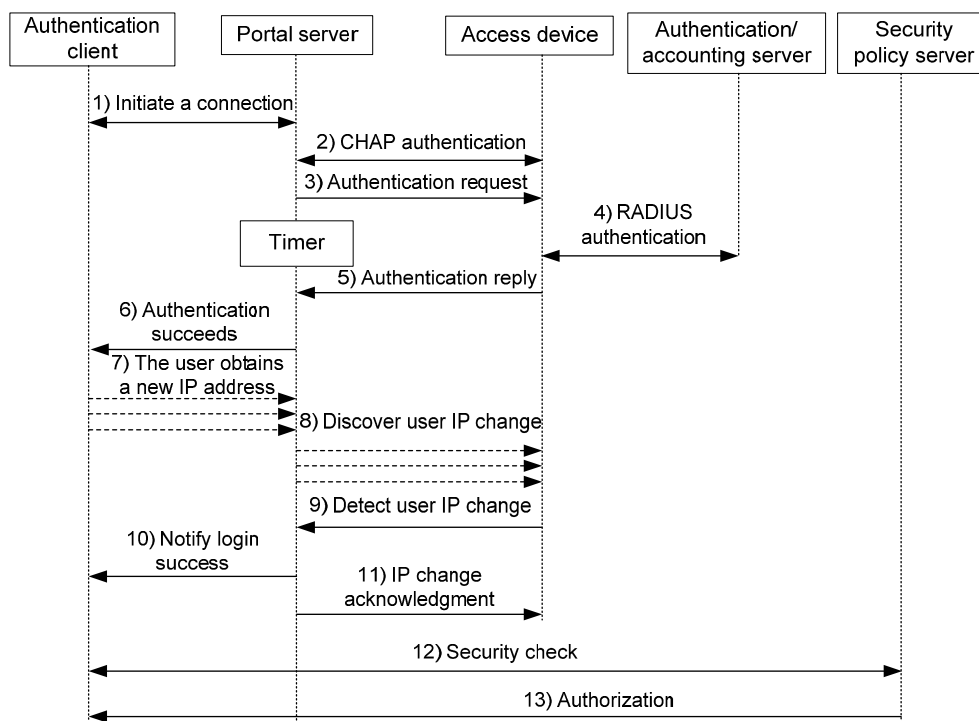
3. The portal server assembles the username and password into an authentication request message and sends it to the access device. Meanwhile, the portal server starts a timer to wait for an authentication acknowledgment message.
4. The access device and the RADIUS server exchange RADIUS packets to authenticate the user.
5. The access device sends an authentication reply to the portal server.
6. The portal server sends an authentication success message to the authentication client to notify it of logon success.
7. The portal server sends an authentication reply acknowledgment message to the access device.

With extended portal functions, the process includes additional steps:

8. The security policy server exchanges security check information with the authentication client to check whether the authentication client meets the security requirements.
9. Based on the security check result, the security policy server authorizes the user to access certain resources, and sends the authorization information to the access device. The access device then controls access of the user based on the authorization information.

Re-DHCP authentication process (with CHAP/PAP authentication)

Figure 56 Re-DHCP authentication process



The re-DHCP authentication takes the following procedure:

The first steps are the same as those in the direct authentication/cross-subnet authentication process.

7. After receiving the authentication success message, the authentication client obtains a new public IP address through DHCP and notifies the portal server that it has obtained a public IP address.
8. The portal server notifies the access device that the authentication client has obtained a new public IP address.
9. Detecting the change of the IP address by examining ARP packets received, the access device notifies the portal server of the change.

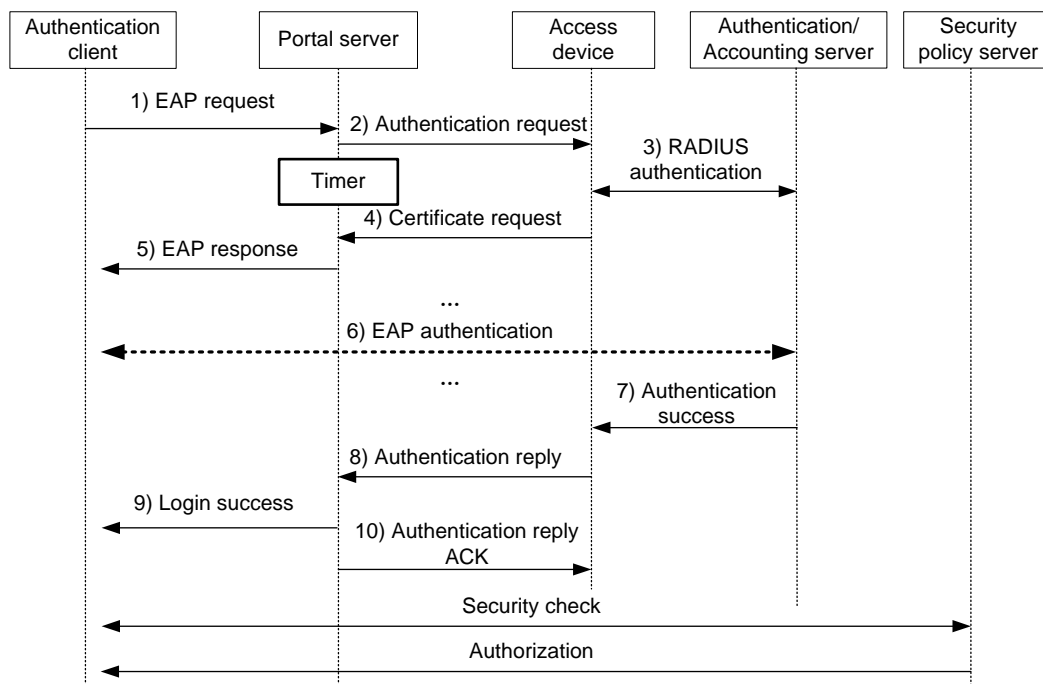
10. The portal server notifies the authentication client of logon success.
11. The portal server sends a user IP address change acknowledgment message to the access device.

With extended portal functions, the process includes additional steps:

12. The security policy server exchanges security check information with the authentication client to check whether the authentication client meets the security requirements.
13. Based on the security check result, the security policy server authorizes the user to access certain resources, and sends the authorization information to the access device. The access device then controls access of the user based on the authorization information.

Portal support for EAP authentication process

Figure 57 Portal support for EAP authentication process



All portal authentication modes share the same EAP authentication steps. The following takes the direct portal authentication as an example to show the EAP authentication process:

1. The authentication client sends an EAP Request/Identity message to the portal server to initiate an EAP authentication process.
2. The portal server sends a portal authentication request to the access device, and starts a timer to wait for the portal authentication reply. The portal authentication request contains several EAP-Message attributes, which are used to encapsulate the EAP packet sent from the authentication client and carry the certificate information of the client.
3. After the access device receives the portal authentication request, it constructs a RADIUS authentication request and sends it to the RADIUS server. The EAP-Message attributes in the RADIUS authentication request are those carried in the received portal authentication request.
4. The access device sends a certificate request to the portal server according to the reply received from the RADIUS server. The certificate request also contains several EAP-Message attributes, which are used to transfer the certificate information of the RADIUS server. The EAP-Message attributes in the certificate request are those carried in the RADIUS authentication reply.

5. After receiving the certificate request, the portal server sends an EAP authentication reply to the authentication client, carrying the EAP-Message attribute values.
6. The authentication client sends another EAP request to continue the EAP authentication with the RADIUS server, during which there may be several portal authentication requests. The subsequent authentication processes are the same as that initiated by the first EAP request, except that the EAP request types vary with the EAP authentication phases.
7. After the authentication client passes the EAP authentication, the RADIUS server sends an authentication reply to the access device. This reply carries the EAP-Success message in the EAP-Message attribute.
8. The access device sends an authentication reply to the portal server. This reply carries the EAP-Success message in the EAP-Message attribute.
9. The portal server notifies the authentication client of the authentication success.
10. The portal server sends an authentication reply acknowledgment to the access device.

The remaining steps are for extended portal authentication. For more information about the steps, see the portal authentication process with CHAP/PAP authentication.

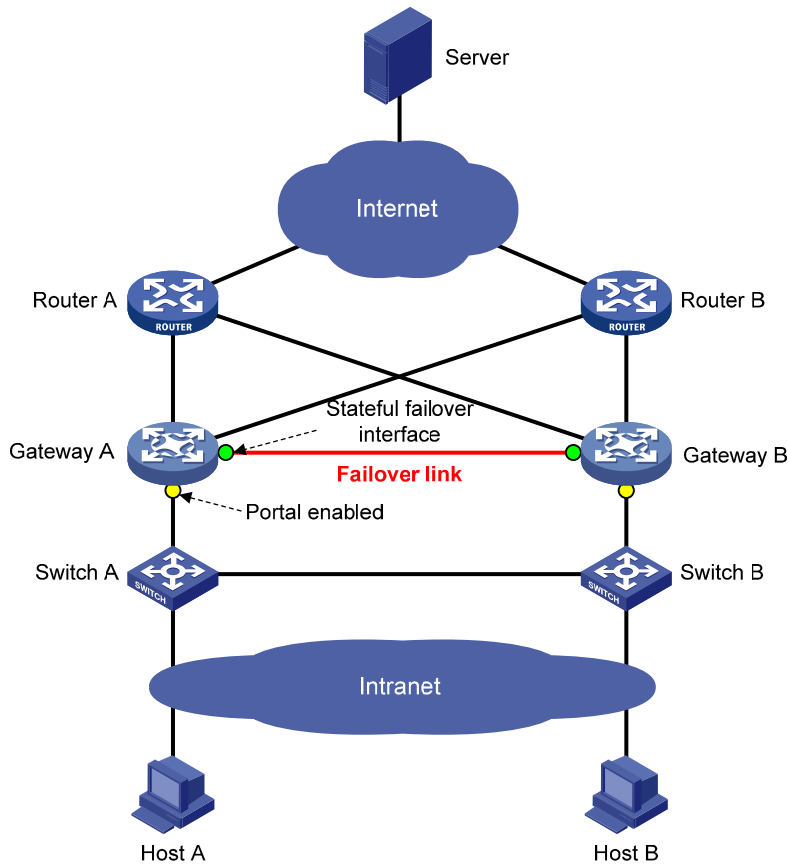
Portal stateful failover (available only on the HP 5500 EI series)

Overview

The stateful failover feature supports hot backup of services on two devices. It can be configured on key devices to avoid service interruptions caused by single point failures. When working normally, the two devices synchronize the service information of each other. If one device fails, the other device takes over the services.

To implement stateful failover, specify a dedicated VLAN (called the "backup VLAN") on each device for stateful failover packets. If both a failover link and a backup VLAN are configured, add the physical ports at the two ends of the failover link to the backup VLAN. For more information about the stateful failover feature, see *High Availability Configuration Guide*.

Figure 58 Network diagram for portal stateful failover configuration



As shown in [Figure 58](#), users have to pass portal authentication to access the Internet. To avoid portal service interruption caused by single point failures, you can deploy two access devices (Gateway A and Gateway B) and configure the portal stateful failover function on them, so that they back up the portal online user information of each other through the failover link. When one of them (Gateway A or Gateway B) fails, the other can guarantee the normal data communication of the online portal users and perform portal authentication for new portal users.

Basic concepts

1. Device states
 - Independence: A stable running status of a device when it does not establish the failover link with the other device.
 - Synchronization: A stable running status of a device when it establishes the failover link with the other device successfully and is ready for data backup.
2. User modes
 - Stand-alone: Indicates that the user data is stored on the local device only. Currently, the local device is in independence state or it is in synchronization state but has not synchronized the user data to the peer device yet.
 - Primary: Indicates that the user logs in from the local device, and the user data is generated on the local device. The local device is in synchronization state and ready for receiving and processing packets from the server.

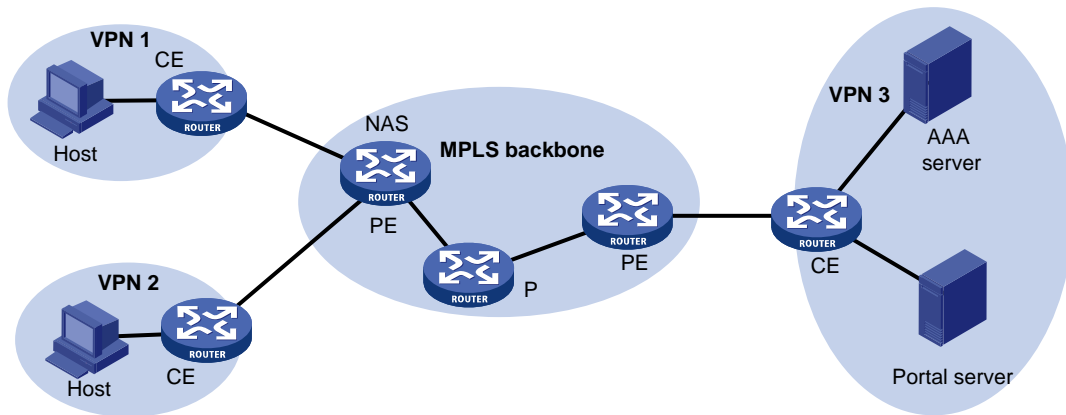
- Secondary: Indicates that the user logs in from the peer device, and the user data is synchronized from the peer device to the local device. The local device is in synchronization state. It only receives and processes the synchronization messages and does not process packets from the server.

Portal authentication across VPNs (available only on the HP 5500 EI series)

This feature is not applicable to VPNs with overlapping address spaces.

In a scenario where the branches belong to different VPNs that are isolated from each other and all portal users in the branches need to be authenticated by the server at the headquarters, you can deploy portal authentication across MPLS VPNs. As shown in [Figure 59](#), the PE connecting the authentication clients serves as the NAS. The NAS is configured with portal authentication and AAA authentication, both of which support authentication across VPNs. The NAS can transmit a client's portal authentication packets in a VPN transparently through the MPLS backbone to the servers in another VPN. This feature implements centralized client authentication across different VPNs while ensuring the separation of packets of the different VPNs.

Figure 59 Network diagram for portal authentication across VPNs



Portal authentication configured on MCE devices can also support authentication across VPNs. For information about MCE, see "[Layer 3—IP Routing Configuration Guid.](#)"

For information about AAA implementation across VPNs, see "[Configuring AAA.](#)"

Portal configuration task list

The HP 5500 SI switch series supports only Layer 2 portal configuration.

Complete these tasks to configure Layer 2 portal authentication:

Task	Remarks	
Specifying the local portal server for Layer 2 portal authentication	Required	
Configuring the local portal server	Customizing authentication pages	Optional
	Configuring the local portal server	Required
Enabling Layer 2 portal authentication	Required	
Controlling access of portal	Configuring a portal-free rule	Optional

Task	Remarks
users	Setting the maximum number of online portal users
	Specifying an authentication domain for portal users
	Configuring Layer 2 portal authentication to support web proxy
	Enabling support for portal user moving
Specifying an Auth-Fail VLAN for portal authentication	Optional
Specifying an auto redirection URL for authenticated portal users	Optional
Configuring online Layer 2 portal user detection	Optional
Logging off portal users	Optional

Complete these tasks to configure Layer 3 portal authentication:

Task	Remarks
Specifying a portal server for Layer 3 portal authentication	Required
Enabling Layer 3 portal authentication	Required
Controlling access of portal users	Configuring a portal-free rule
	Configuring an authentication source subnet
	Setting the maximum number of online portal users
	Specifying an authentication domain for portal users
Configuring RADIUS related attributes	Specifying NAS-Port-Type for an interface
	Specifying a NAS ID profile for an interface
Specifying a source IP address for outgoing portal packets	Optional
Configuring portal stateful failover (available only on the HP 5500 EI series)	Optional
Specifying an auto redirection URL for authenticated portal users	Optional
Configuring portal detection functions	Configuring the portal server detection function
	Configuring portal user information synchronization
Logging off portal users	Optional

Configuration prerequisites

The portal feature provides a solution for user identity authentication and security check. However, the portal feature cannot implement this solution by itself. RADIUS authentication needs to be configured on the access device to cooperate with the portal feature to complete user authentication.

The prerequisites for portal authentication configuration are as follows:

- The portal server and the RADIUS server have been installed and configured properly. Local portal authentication requires no independent portal server be installed.

- With re-DHCP authentication, the IP address check function of the DHCP relay agent is enabled on the access device, and the DHCP server is installed and configured properly. (Available only on the HP 5500 EI series)
- The portal client, access device, and servers can reach each other.
- With RADIUS authentication, usernames and passwords of the users are configured on the RADIUS server, and the RADIUS client configurations are performed on the access device. For information about RADIUS client configuration, see "[Configuring AAA.](#)"
- To implement extended portal functions, install and configure IMC EAD, and make sure that the ACLs configured on the access device correspond to those specified for the resources in the quarantined area and for the restricted resources on the security policy server. For information about security policy server configuration on the access device, see "[Configuring AAA.](#)"

For installation and configuration about the security policy server, see *IMC EAD Security Policy Help*.

The ACL for resources in the quarantined area and that for restricted resources correspond to isolation ACL and security ACL, respectively, on the security policy server.

You can modify the authorized ACLs on the access device. However, your changes take effect only for portal users logging on after the modification.

For portal authentication to work normally, make sure that the system name of the access device is no more than 16 characters.

Specifying the portal server

Specifying the local portal server for Layer 2 portal authentication

Layer 2 portal authentication uses the local portal server. Specify the IP address of a Layer 3 interface on the device that is routable to the portal client as the listening IP address of the local portal server. HP recommends using the IP address of a loopback interface rather than a physical Layer 3 interface, because:

- The status of a loopback interface is stable. There will be no authentication page access failures caused by interface failures.
- A loopback interface does not forward received packets to any network, avoiding impact on system performance when there are many network access requests.

To specify the local portal server for Layer 2 portal authentication:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Specify the listening IP address of the local portal server for Layer 2 portal authentication.	portal local-server ip <i>ip-address</i>	By default, no listening IP address is specified.

NOTE:

The specified listening IP address can be changed or deleted only if Layer 2 portal authentication is not enabled on any port.

Specifying a portal server for Layer 3 portal authentication (available only on the HP 5500 EI series)

This task allows you to specify the portal server parameters for Layer 3 portal authentication, including the portal server IP address, shared encryption key, server port, and the URL address for web authentication. According to the networking environment, you can configure a remote portal server or a local portal server as needed.

- To configure a remote portal server, specify the IP address of the remote portal server.

When you specify a portal server for Layer 3 portal authentication, follow these guidelines:

- If the portal server is in an MPLS VPN, specify the VPN instance when specifying the portal server on the device, so the device can send packets to the portal server.

To specify a portal server for Layer 3 authentication:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Specify a portal server and configure related parameters.	portal server <i>server-name</i> ip <i>ip-address</i> [key [cipher simple] <i>key-string</i> port <i>port-id</i> url <i>url-string</i> vpn-instance <i>vpn-instance-name</i>] * ipv6 <i>ipv6-address</i> [key [cipher simple] <i>key-string</i> port <i>port-id</i> url <i>url-string</i>] * }	By default, no portal server is specified.

NOTE:

The specified parameters of a portal server can be modified or deleted only if the portal server is not referenced on any interface.

Configuring the local portal server

Configuring a local portal server is required only for local portal authentication. During local portal authentication, the local portal server pushes authentication pages to users. You can define the authentication pages for users; otherwise, the default authentication pages will be used during the authentication process.

Customizing authentication pages

Customized authentication pages exist in the form of HTML files. You can compress them and then save them in the storage medium of the access device.

A set of authentication pages includes six main authentication pages and their page elements. The six main authentication pages are the logon page, the logon success page, the logon failure page, the online page, the system busy page, and the logoff success page. The page elements refer to the files that the authentication pages reference, for example, **back.jpg** for page **Logon.htm**. Each main authentication page can reference multiple page elements. If you define only some of the main authentication pages, the system will use the default authentication pages for the undefined ones.

For the local portal server to operate normally and steadily, follow the following rules when customizing authentication pages:

Rules on file names

The main authentication pages have predefined file names, which cannot be changed.

Table 10 Main authentication page file names

Main authentication page	File name
Logon page	logon.htm
Logon success page	logonSuccess.htm
Logon failure page	logonFail.htm
Online page Pushed after the user gets online for online notification	online.htm
System busy page Pushed when the system is busy or the user is in the logon process	busy.htm
Logoff success page	logoffSuccess.htm

NOTE:

You can define the names of the files other than the main authentication page files. The file names and directory names are case-insensitive.

Rules on page requests

The local portal server supports only Post and Get requests.

- Get requests are used to get the static files in the authentication pages and allow no recursion. For example, if file Logon.htm includes contents that perform Get action on file ca.htm, file ca.htm cannot include any reference to file Logon.htm.
- Post requests are used when users submit username and password pairs, log on the system, and log off the system.

Rules on Post request attributes

1. Observe the following requirements when editing a form of an authentication page:
 - An authentication page can have multiple forms, but there must be one and only one form whose action is **logon.cgi**. Otherwise, user information cannot be sent to the local portal server.
 - The username attribute is fixed as **PtUser**, and the password attribute is fixed as **PtPwd**.
 - Attribute **PtButton** is required to indicate the action that the user requests, which can be **Logon** or **Logoff**.
 - A logon Post request must contain **PtUser**, **PtPwd**, and **PtButton** attributes.
 - A logoff Post request must contain the **PtButton** attribute.
2. Authentication pages **logon.htm** and **logonFail.htm** must contain the logon Post request.

The following example shows part of the script in page **logon.htm**.

```
<form action=logon.cgi method = post >  
<p>User name:<input type="text" name = "PtUser" style="width:160px;height:22px"  
maxlength=64>
```

```

<p>Password :<input type="password" name = "PtPwd" style="width:160px;height:22px"
maxlength=32>
<p><input type=SUBMIT value="Logon" name = "PtButton" style="width:60px;"
onclick="form.action=form.action+location.search;"
</form>

```

3. Authentication pages **logonSuccess.htm** and **online.htm** must contain the logoff Post request.

The following example shows part of the script in page **online.htm**.

```

<form action=logon.cgi method = post >
<p><input type=SUBMIT value="Logoff" name="PtButton" style="width:60px;" >
</form>

```

Rules on page file compression and saving

- A set of authentication page files must be compressed into a standard zip file. The name of a zip file can contain only letters, numerals, and underscores. The zip file of the default authentication pages must be saved with name **defaultfile.zip**.
- The set of authentication pages must be located in the root directory of the zip file.
- Zip files can be transferred to the device through FTP or TFTP. The default authentication pages file must be saved in the root directory of the device, and other authentication files can be saved in the root directory or the **portal** directory under the root directory of the device.

Examples of zip files on the device:

```

<Sysname> dir
Directory of flash:/portal/
 0  -rw-    1405  Feb 28 2011 15:53:31  ssid2.zip
 1  -rw-    1405  Feb 28 2011 15:53:20  ssid1.zip
 2  -rw-    1405  Feb 28 2011 15:53:39  ssid3.zip
 3  -rw-    1405  Feb 28 2011 15:53:44  ssid4.zip
2540 KB total (1319 KB free)

```

Rules on file size and contents

For the system to push customized authentication pages smoothly, you need comply with the following size and content requirements on authentication pages.

- The size of the zip file of each set of authentication pages, including the main authentication pages and the page elements, must be no more than 500 KB.
- The size of a single page, including the main authentication page and its page elements, must be no more than 50 KB before being compressed.
- Page elements can contain only static contents such as HTML, JS, CSS, and pictures.

Logging off a user who closes the logon success or online page

After a user passes authentication, the system pushes the logon success page named logonSuccess.htm. If the user initiates another authentication through the logon page, the system pushes the online page named online.htm. You can configure the device to forcibly log off the user when the user closes either of these two pages. To do so, add the following contents in logonSuccess.htm and online.htm:

1. Reference to JS file pt_private.js.
2. Function pt_unload(), which is used to trigger page unloading.
3. Function pt_submit(), the event handler function for Form.
4. Function pt_init(), which is for triggering page loading.

The following is a script example with the added contents highlighted in gray:

```

<html>
<head>
<script type="text/javascript" language="javascript" src="pt_private.js"></script>
</head>
<body onload="pt_init();" onbeforeunload="return pt_unload();">
... ..
<form action=logon.cgi method = post onsubmit="pt_submit()">
... ..
</body>
</html>

```

Redirecting authenticated users to a specified web page

To make the device automatically redirect authenticated users to a specified web page, do the following in logon.htm and logonSuccess.htm:

1. In logon.htm, set the target attribute of Form to **blank**.

See the contents in gray:

```
<form method=post action=logon.cgi target="blank">
```

2. Add the function for page loading pt_init() to logonSuccess.htm.

See the contents in gray:

```

<html>
<head>
<title>LogonSucceeded</title>
<script type="text/javascript" language="javascript"
src="pt_private.js"></script>
</head>
<body onload="pt_init();" onbeforeunload="return pt_unload();">
... ..
</body>
</html>

```

HP recommends using browser IE 6.0 or above on the authentication clients. Make sure the browser of an authentication client permits pop-ups or permits pop-ups from the access device. Otherwise, the user cannot log off by closing the logon success or online page and can only click **Cancel** to return back to the logon success or online page.

If a user refreshes the logon success or online page, or jumps to another web site from either of the pages, the device also logs off the user.

Only IE, Firefox, and Safari browsers support the device to log off the user when the user closes the logon success or online page. Other browsers, such as Chrome and Opera do not support this function.

Configuring the local portal server

To make the local portal server take effect, specify the protocol to be used for communication between the portal client and local portal server.

Configuration prerequisites

To configure the local portal server to support HTTPS, complete these configurations at first:

- Configure PKI policies, obtain the CA certificate, and apply for a local certificate. For more information, see "[Configuring PKI](#)."
- Configure the SSL server policy, and specify the PKI domain to be used, which is configured in the above step. For more information, see "[Configuring SSL](#)."

When you specify the protocol for the local portal server to support, the local portal server will load the default authentication page file, which is supposed to be saved in the root directory of the device. Therefore, to make sure that the local portal server uses the user-defined default authentication pages, you must edit and save them properly. Otherwise, the system default authentication pages are used.

Configuration procedure

To configure the local portal server:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the protocol type for the local portal server to support and load the default authentication page file.	portal local-server { http https server-policy <i>policy-name</i> }	By default, the local portal server does not support any protocol.
3. Configure the welcome banner of the default authentication pages of the local portal server.	portal server banner <i>banner-string</i>	Optional. No welcome banner by default.

Enabling portal authentication

Only after you enable portal authentication on an access interface, can the access interface perform portal authentication for connected clients.

Enabling Layer 2 portal authentication

Before enabling Layer 2 portal authentication, make sure that:

- The listening IP address of the local portal server is specified.
- Layer 3 portal authentication is not enabled on any interface.

Follow these guidelines when you enable Layer 2 portal authentication:

- To ensure normal operation of portal authentication on a Layer 2 port, do not enable port security, guest VLAN of 802.1X, or EAD fast deployment of 802.1X on the port.
- To support assignment of authorized VLANs, you must enable the MAC-based VLAN function on the port.

To enable Layer 2 portal authentication:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet port view.	interface <i>interface-type</i> <i>interface-number</i>	N/A

Step	Command	Remarks
3.	Enable Layer 2 portal authentication on the port.	portal local-server enable Not enabled by default.

Enabling Layer 3 portal authentication (available only on the HP 5500 EI series)

Before enabling Layer 3 portal authentication on an interface, make sure that:

- An IP address is configured for the interface.
- The interface is not added to any port aggregation group.
- Layer 2 portal authentication is not enabled on any ports.

Follow these guidelines when you enable Layer 3 portal authentication:

- The destination port number that the device uses for sending unsolicited packets to the portal server must be the same as the port number that the remote portal server actually uses.
- Cross-subnet authentication mode (**portal server server-name method layer3**) does not require Layer 3 forwarding devices between the access device and the authentication clients. However, if Layer 3 forwarding devices exist between the authentication client and the access device, you must select the cross-subnet portal authentication mode.
- In re-DHCP authentication mode, a client can use a public IP address to send packets before passing portal authentication. However, responses to the packets are restricted.
- An IPv6 portal server does not support the re-DHCP portal authentication mode.
- You can enable both an IPv4 portal server and an IPv6 portal server for Layer 3 portal authentication on an interface, but you cannot enable two IPv4 or two IPv6 portal servers on the interface.

To enable Layer 3 portal authentication:

Step	Command	Remarks
1.	Enter system view.	system-view N/A
2.	Enter interface view.	interface <i>interface-type</i> <i>interface-number</i> N/A
3.	Enable Layer 3 portal authentication on the interface.	portal server server-name method { direct layer3 redhcp } Not enabled by default.

NOTE:

The portal server and its parameters can be deleted or modified only when the portal server is not referenced by any interface.

Controlling access of portal users

Configuring a portal-free rule

A portal-free rule allows specified users to access specified external websites without portal authentication.

The matching items for a portal-free rule include the source and destination IP address, source MAC address, inbound interface, and VLAN. Packets matching a portal-free rule will not trigger portal authentication, so that users sending the packets can directly access the specified external websites.

For Layer 2 portal authentication, you can configure only a portal-free rule that is from any source address to any or a specified destination address. If you configure a portal-free rule that is from any source address to a specified destination address, users can access the specified address directly, without being redirected to the portal authentication page for portal authentication. Usually, you can configure the IP address of a server that provides certain services (such as software upgrading service) as the destination IP address of a portal-free rule, so that Layer 2 portal authentication users can access the services without portal authentication.

Follow these guidelines when you configure a portal-free rule:

- If you specify both a VLAN and an interface in a portal-free rule, the interface must belong to the VLAN. Otherwise, the rule does not take effect.
- You cannot configure two or more portal-free rules with the same filtering criteria. Otherwise, the system prompts that the rule already exists.
- A Layer 2 interface in an aggregation group cannot be specified as the source interface of a portal-free rule, and the source interface of a portal-free rule cannot be added to an aggregation group.

To configure a portal-free rule:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure a portal-free rule.	<ul style="list-style-type: none">• To configure an IPv4 portal-free rule: portal free-rule <i>rule-number</i> { destination { any ip { <i>ip-address</i> mask { <i>mask-length</i> <i>netmask</i> } any } } source { any [interface <i>interface-type interface-number</i> ip { <i>ip-address</i> mask { <i>mask-length</i> <i>mask</i> } any } mac <i>mac-address</i> vlan <i>vlan-id</i>] * } } *• To configure an IPv6 portal-free rule: portal free-rule <i>rule-number</i> { destination { any ipv6 { <i>ipv6-address</i> <i>prefix-length</i> any } } source { any [interface <i>interface-type</i> <i>interface-number</i> ipv6 { <i>ipv6-address</i> <i>prefix-length</i> any } mac <i>mac-address</i> vlan <i>vlan-id</i>] * } } *	Configure at least one command.

NOTE:

Regardless of whether portal authentication is enabled or not, you can only add or remove a portal-free rule. You cannot modify it.

Configuring an authentication source subnet (available only on the HP 5500 EI series)

Only Layer 3 portal authentication supports this feature.

By configuring authentication source subnets, you specify that only HTTP packets from users on the authentication source subnets can trigger portal authentication. If an unauthenticated user is not on any authentication source subnet, the access device discards all the user's HTTP packets that do not match any portal-free rule.

To configure an authentication source subnet:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure an authentication source subnet.	portal auth-network <i>network-address</i> { <i>mask-length</i> <i>mask</i> }	Optional. By default, the source IPv4 subnet is 0.0.0.0/0, and the source IPv6 subnet is ::/0, meaning that users from any IPv4 or IPv6 subnet must pass portal authentication to access network resources. You can configure multiple authentication source subnets by executing the portal auth-network command repeatedly.

NOTE:

Configuration of authentication source subnets applies to only cross-subnet authentication. In direct authentication mode, the authentication source subnet is 0.0.0.0/0. In re-DHCP authentication mode, the authentication source subnet of an interface is the subnet to which the private IP address of the interface belongs.

Setting the maximum number of online portal users

You can use this feature to control the total number of online portal users in the system.

If the maximum number of online portal users to be set is less than that of the current online portal users, the limit can be set successfully and does not impact the online portal users, but the system does not allow new portal users to log on until the number drops down below the limit.

To set the maximum number of online portal users allowed in the system:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Set the maximum number of online portal users.	portal max-user <i>max-number</i>	By default, the maximum number is 3000 on the HP 5500 EI series and 1000 on the HP 5500 SI series.

NOTE:

The maximum number of online portal users the switch actually assigns depends on the ACL resources on the switch.

Specifying an authentication domain for portal users

After you specify an authentication domain for portal users on an interface, the device uses the authentication domain for authentication, authorization, and accounting (AAA) of all portal users on the interface, ignoring the domain names carried in the usernames. This allows you to specify different authentication domains for different interfaces as needed.

To specify an authentication domain for portal users on an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Specify an authentication domain for portal users on the interface.	portal domain [<i>ipv6</i>] <i>domain-name</i>	By default, no authentication domain is specified for portal users.

The switch selects the authentication domain for a portal user on an interface in this order: the authentication domain specified for the interface, the authentication domain carried in the username, and the system default authentication domain. For information about the default authentication domain, see "[Configuring AAA](#)."

Configuring Layer 2 portal authentication to support web proxy

By default, proxied HTTP requests cannot trigger Layer 2 portal authentication but are silently dropped. To allow such HTTP requests to trigger portal authentication, configure the port numbers of the web proxy servers on the switch.

If a user's browser uses the Web Proxy Auto-Discovery (WPAD) protocol to discover web proxy servers, add the port numbers of the web proxy servers on the switch, and configure portal-free rules to allow user packets destined for the IP address of the WPAD server to pass without authentication.

You must add the port numbers of the web proxy servers on the switch and users must make sure their browsers that use a web proxy server do not use the proxy server for the listening IP address of the local portal server. Thus, HTTP packets that the portal user sends to the local portal server are not sent to the web proxy server.

To configure Layer 2 portal authentication to support a web proxy:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Add a web proxy server port number.	portal web-proxy port <i>port-number</i>	By default, no web proxy server port number is configured and proxied HTTP requests cannot trigger portal authentication.

Enabling support for portal user moving

Only Layer 2 portal authentication supports this feature.

In scenarios where there are hubs, Layer 2 switches, or APs between users and the access devices, if an authenticated user moves from the current access port to another Layer 2-portal-authentication-enabled port of the device without logging off, the user cannot get online when the original port is still up. The reason is that the original port is still maintaining the authentication information of the user and the device does not permit such a user to get online from another port by default.

To solve the problem described above, enable support for portal user moving on the device. Then, when a user moves from a port of the device to another, the device provides services in either of the following ways:

- If the original port is still up and the two ports belong to the same VLAN, the device allows the user to continue to access the network without re-authentication, and uses the new port information for user accounting.
- If the original port is down or the two ports belong to different VLANs, the device removes the authentication information of the user from the original port and authenticates the user on the new port.

To enable support for portal user moving:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enable support for portal user moving.	portal move-mode auto	Disabled by default

For a user with authorization information (such as authorized VLAN) configured, after the user moves from a port to another, the switch tries to assign the authorization information to the new port. If the operation fails, the switch deletes the user's information from the original port and re-authenticates the user on the new port.

Specifying an Auth-Fail VLAN for portal authentication

Only Layer 2 portal authentication supports this feature.

This task sets the Auth-Fail VLAN to be assigned to users failing portal authentication. You can specify different Auth-Fail VLANs for portal authentication on different ports. A port can be specified with only one Auth-Fail VLAN for portal authentication.

Before specifying an Auth-Fail VLAN, be sure to create the VLAN.

To specify an Auth-Fail VLAN for portal authentication:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter Layer 2 Ethernet interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Specify an Auth-Fail VLAN for portal authentication on the port.	portal auth-fail vlan <i>authfail-vlan-id</i>	Not specified by default

After you specify an Auth-Fail VLAN for portal authentication on a port, you must also enable the MAC-based VLAN function on the port to make the specified Auth-Fail VLAN take effect. For information about MAC VLAN, see *Layer 2—LAN Switching Configuration Guide*.

The MAC-VLAN entries generated in response to portal authentication failures do not overwrite the MAC-VLAN entries already generated in other authentication modes.

Configuring RADIUS related attributes

Only Layer 3 portal authentication supports this feature.

Specifying NAS-Port-Type for an interface

NAS-Port-Type is a standard RADIUS attribute for indicating a user access port type. With this attribute specified on an interface, when a portal user logs on from the interface, the device uses the specified NAS-Port-Type value as that in the RADIUS request to be sent to the RADIUS server. If NAS-Port-Type is not specified, the device uses the access port type obtained.

If there are multiple network devices between the Broadband Access Server (BAS, the portal authentication access device) and a portal client, the BAS may not be able to obtain a user's correct access port information. For example, for a wireless client using portal authentication, the access port type obtained by the BAS may be the type of the wired port that authenticates the user. To make sure that the BAS delivers the right access port information to the RADIUS server, specify the NAS-Port-Type according to the practical access environment.

To specify the NAS-Port-Type value for an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Specify the NAS-Port-Type value for the interface.	portal nas-port-type { ethernet wireless }	Not configured by default

Specifying a NAS ID profile for an interface

In some networks, users' access points are identified by their access VLANs. Network carriers need to use NAS-identifiers to identify user access points. With a NAS ID profile specified on an interface, when a user logs in from the interface, the access device checks the specified profile to obtain the NAS ID that is bound with the access VLAN. The value of this NAS ID is used as that of the NAS-identifier attribute in the RADIUS packets to be sent to the RADIUS server.

A NAS ID profile defines the binding relationship between VLANs and NAS IDs. A NAS ID-VLAN binding is defined by the **nas-id id-value bind vlan vlan-id** command, which is described in detail in AAA configuration commands in the *Security Command Reference*.

If no NAS-ID profile is specified for an interface or no matching binding is found in the specified profile, the switch uses the device name as the interface NAS ID.

To configure a NAS ID profile for an interface:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Create a NAS ID profile and enter NAS ID profile view.	aaa nas-id profile <i>profile-name</i>	For more information about the command, see <i>Security Command Reference</i> .
3. Bind a NAS ID with a VLAN.	nas-id nas-identifier bind vlan <i>vlan-id</i>	For more information about the command, see <i>Security Command Reference</i> .
4. Return to system view.	quit	N/A
5. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
6. Specify a NAS ID profile for the interface.	portal nas-id-profile <i>profile-name</i>	By default, an interface is specified with no NAS ID profile.

Specifying a source IP address for outgoing portal packets

After you specify a source IP address for outgoing portal packets on an interface, the IP address is used as the source IP address of packets that the access device sends to the portal server, and the destination IP address of packets that the portal server sends to the access device.

To specify a source IP address for outgoing portal packets to be sent:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A

Step	Command	Remarks
3.	Specify a source IP address for outgoing portal packets.	Optional. By default, no source IP address is specified for outgoing portal packets and the IP address of the user logon interface is used as the source IP address of the outgoing portal packets. In NAT environments, HP recommends specifying the interface's public IP address as the source IP address of outgoing portal packets.

Configuring portal stateful failover (available only on the HP 5500 E1 series)

Only Layer 3 portal authentication supports this feature.

To implement stateful failover for portal, configure VRRP for traffic switchover, and perform the following configurations for service backup on each of the two devices that back up each other:

- Specify an interface for backing up portal services, which is called portal service backup interface in this document, and enable portal on the portal service backup interface. The portal service backup interface is different from the stateful failover interface. Stateful failover interfaces only forward state negotiation messages and backup data.
- Specify the portal group to which the portal service backup interface belongs. Be sure to specify the same portal group for the portal service backup interfaces that back up each other on the two devices.
- Specify the device ID. Make sure that the device ID of the local device is different from that of the peer device.
- Specify the backup source IP address for RADIUS packets to be sent as the source IP address for RADIUS packets that is configured on the peer device, so that the peer device can receive packets from the server. (This configuration is optional.)
- Specify the backup VLAN, and enable stateful failover. For related configuration, see *High Availability Configuration Guide*.

After the working state of the two devices changes from independence to synchronization and the portal group takes effect, the two devices start to back up the data of online portal users for each other.

The AAA and portal configuration must be consistent on the two devices that back up each other. For example, you must configure the same portal server on the two devices.

To configure stateful failover:

Step	Command	Remarks
1.	Enter system view.	N/A
2.	Enter interface view.	N/A

Step	Command	Remarks
3. Specify the portal group to which the portal service backup interface belongs.	portal backup-group <i>group-id</i>	By default, the portal service backup interface does not belong to any portal group. The portal service backup interfaces on the two devices for stateful failover must belong to the same portal group.
4. Return to system view.	quit	N/A
5. Specify the device ID in stateful failover mode.	nas device-id <i>device-id</i>	By default, the device operates in stand-alone mode, and thus has no device ID configured. For more information about the command, see <i>Security Command Reference</i> .
	radius nas-backup-ip <i>ip-address</i>	Optional. Use either approach. By default, no backup source IP address is specified.
6. Specify the backup source IP address for RADIUS packets to be sent.	radius scheme <i>radius-scheme-name</i> nas-backup-ip <i>ip-address</i>	You do not need to specify the backup source IP address if the device uses the virtual IP address of the VRRP group to which the uplink belongs as the source IP address of outgoing RADIUS packets. For more information about the command, see <i>Security Command Reference</i> .

After you configure portal stateful failover for two devices, note the following issues:

- In stateful failover mode, the device does not support re-DHCP portal authentication on the portal service backup interface.
- In stateful failover mode, if a user on either device is logged out, the information of the user on the other device is deleted, too. You can log off a user on the device or on the portal server. For example, you can use the **cut connection** and **portal delete-user** commands on the device to log off users.
- Specifying or changing the device ID of a device will log off all online users on the device. Therefore, perform the configuration only when necessary and, after the configuration, save the configuration and restart the device.
- Do not delete the configured backup source IP addresses. Otherwise, online users on the backup device may not be able to receive packets from the server.

Specifying an auto redirection URL for authenticated portal users

After a user passes portal authentication, if the access device is configured with an auto redirection URL, it redirects the user to the URL after a specified period of time.

Follow these guidelines to specify an auto redirection URL for authenticated portal users:

- To use this feature for remote Layer 3 portal authentication, the portal server must be the IMC portal server that supports the page auto-redirection function.
- The **wait-time** *period* option is effective to only local portal authentication.
- When no auto redirection URL is specified for authenticated portal users, an authenticated user is usually redirected to the URL the user typed in the address bar before portal authentication. However, with local portal authentication, if the URL a user typed in the address bar before portal authentication is more than 255 characters, the user cannot be redirected to the page of the URL after passing portal authentication.

To specify an auto redirection URL for authenticated portal users:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Specify an auto redirection URL for authenticated portal users.	portal redirect-url <i>url-string</i> [wait-time <i>period</i>]	By default, an authenticated user is redirected to the URL the user typed in the address bar before portal authentication.

Configuring portal detection functions

Configuring online Layer 2 portal user detection

Only Layer 2 portal authentication supports this feature.

After a Layer 2 portal user gets online, the device starts a detection timer for the user, and checks whether the user's MAC address entry has been aged out or the user's MAC address entry has been matched (a match means a packet has been received from the user) at the interval. If the device finds no MAC address entry for the user or receives no packets from the user during two successive detection intervals, the device considers that the user has gone offline and clears the authentication information of the user.

To set the Layer 2 portal user detection interval:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Enter interface view.	interface <i>interface-type</i> <i>interface-number</i>	N/A
3. Set the Layer 2 portal user detection interval.	portal offline-detect interval <i>offline-detect-interval</i>	300 seconds by default

Configuring the portal server detection function (available only on the HP 5500 EI series)

Only Layer 3 portal authentication supports this feature.

During portal authentication, if the communication between the access device and portal server is broken, new portal users are not able to log on and the online portal users are not able to log off normally. To address this problem, the access device must be able to detect the reachability changes of the portal server quickly and take corresponding actions to deal with the changes. For example, after the access device detects that the portal server is unreachable, it allows portal users to access network resources without authentication. This function is referred to as "portal authentication bypass." It allows for flexible user access control.

With the portal server detection function, the device can detect the status of a specific portal server. The specific configurations include:

1. Detection methods (you can choose either or both)
 - **Probing HTTP connections**—The access device periodically sends TCP connection requests to the HTTP service port of the portal servers configured on its interfaces. If the TCP connection with a portal server can be established, the access device considers that the probe succeeds (the HTTP service of the portal server is open and the portal server is reachable). If the TCP connection cannot be established, the access device considers that the probe fails and the portal server is unreachable.
 - **Probing portal heartbeat packets**—A portal server that supports the portal heartbeat function, (only the IMC portal server supports this function), sends portal heartbeat packets to portal access devices periodically. If an access device receives a portal heartbeat packet or an authentication packet within a probe interval, the access device considers that the probe succeeds and the portal server is reachable; otherwise, it considers that the probe fails and the portal server is unreachable.
2. Probe parameters
 - **Probe interval**—Interval at which probe attempts are made.
 - **Maximum number of probe attempts**—Maximum number of consecutive probe attempts allowed. If the number of consecutive probes reaches this value, the access device considers that the portal server is unreachable.
3. Actions to be taken when the server reachability status changes (you can choose one or more)
 - **Sending a trap message**—When the status of a portal server changes, the access device sends a trap message to the network management server (NMS). The trap message contains the portal server name and the current state of the portal server.
 - **Sending a log**—When the status of a portal server changes, the access device sends a log message. The log message indicates the portal server name and the current state and original state of the portal server.
 - **Disabling portal authentication (enabling portal authentication bypass)**—When the device detects that a portal server is unreachable, it disables portal authentication on the interfaces that use the portal server (allows all portal users on the interfaces to access network resources). When the device receives from the portal server portal heartbeat packets or authentication packets (such as logon requests and logout requests), it re-enables the portal authentication function.

You can configure any combination of the configuration items described as needed, with respect to the following:

- If both detection methods are specified, a portal server is regarded as unreachable as long as one detection method fails, and an unreachable portal server is regarded as recovered only when both detection methods succeed.
- If multiple actions are specified, the access device executes all the specified actions when the status of a portal server changes.

- The detection function configured for a portal server takes effect on an interface only after you enable portal authentication and reference the portal server on the interface.

To configure the portal server detection function:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the portal server detection function.	portal server <i>server-name</i> server-detect method { http portal-heartbeat } * action { log permit-all trap } * [interval <i>interval</i>] [retry <i>retries</i>]	Not configured by default. The portal server specified in the command must exist.

The portal heartbeat detection method works only when the portal server supports the portal server heartbeat function. Only the IMC portal server supports the portal server heartbeat function. To implement detection with this method, you also need to configure the portal server heartbeat function on the IMC portal server and make sure that the product of **interval** and **retry** is greater than or equal to the portal server heartbeat interval. HP recommends configuring the **interval** to be greater than the portal server heartbeat interval configured on the portal server.

Configuring portal user information synchronization (available only on the HP 5500 EI series)

Only Layer 3 portal authentication supports this feature.

Once the device loses communication with a portal server, the portal user information on the device and that on the portal server may be inconsistent after the communication resumes. To solve this problem, the device provides the portal user information synchronization function. This function is implemented by sending and detecting the portal synchronization packet. The process is as follows:

1. The portal server sends the online user information to the access device in a user synchronization packet at the user heartbeat interval, which is set on the portal server.
2. Upon receiving the user synchronization packet, the access device checks the user information carried in the packet with its own. If the device finds a nonexistent user in the packet, it informs the portal server of the information and the portal server will delete the user. If the device finds that one of its users does not appear in the user synchronization packets within N consecutive synchronization probe intervals (N is equal to the value of *retries* configured in the **portal server user-sync** command), it considers that the user does not exist on the portal server and logs the user off.

To configure the portal user information synchronization function:

Step	Command	Remarks
1. Enter system view.	system-view	N/A
2. Configure the portal user information synchronization function.	portal server <i>server-name</i> user-sync [interval <i>interval</i>] [retry <i>retries</i>]	Not configured by default. The portal server specified in the command must exist. This function can take effect only when the specified portal server is referenced on the interface connecting the users.

The user information synchronization function requires that a portal server supports the portal user heartbeat function. Only the IMC portal server supports the portal user heartbeat function. To implement the portal user synchronization function, you also need to configure the user heartbeat function on the portal server and make sure the product of **interval** and **retry** is greater than or equal to the portal user heartbeat interval. HP recommends that you configure the **interval** to be greater than the portal user heartbeat interval configured on the portal server.

For redundant user information on the device (information for users who are considered nonexistent on the portal server), the device deletes the information during the (N+1)th interval, where N is equal to the value of *retries* configured in the **portal server user-sync** command.

Logging off portal users

Logging off a user terminates the authentication process for the user or removes the user from the authenticated users list.

To log off users:

Step	Command
1. Enter system view.	system-view
2. Log off users.	portal delete-user { <i>ipv4-address</i> all interface <i>interface-type</i> <i>interface-number</i> ipv6 <i>ipv6-address</i> }

Displaying and maintaining portal

Task	Command	Remarks
Display the ACLs on an interface (available only on the HP 5500 EI series).	display portal acl { all dynamic static } interface <i>interface-type</i> <i>interface-number</i> [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display portal connection statistics on a specific interface or all interfaces (available only on the HP 5500 EI series).	display portal connection statistics { all interface <i>interface-type</i> <i>interface-number</i> } [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about a portal-free rule or all portal-free rules.	display portal free-rule [<i>rule-number</i>] [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display the portal configuration of an interface.	display portal interface <i>interface-type</i> <i>interface-number</i> [[{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display configuration information about the local portal server.	display portal local-server [[{ begin exclude include } <i>regular-expression</i>]	Available in any view

Task	Command	Remarks
Display information about a specific portal server or all portal servers (available only on the HP 5500 EI series).	display portal server [<i>server-name</i>] [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display portal server statistics on a specific interface or all interfaces (available only on the HP 5500 EI series).	display portal server statistics { all interface <i>interface-type interface-number</i> } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display TCP spoofing statistics.	display portal tcp-cheat statistics [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Display information about portal users on a specific interface or all interfaces.	display portal user { all interface <i>interface-type interface-number</i> } [{ begin exclude include } <i>regular-expression</i>]	Available in any view
Clear portal connection statistics on a specific interface or all interfaces (available only on the HP 5500 EI series).	reset portal connection statistics { all interface <i>interface-type interface-number</i> }	Available in user view
Clear portal server statistics on a specific interface or all interfaces (available only on the HP 5500 EI series).	reset portal server statistics { all interface <i>interface-type interface-number</i> }	Available in user view
Clear TCP spoofing statistics.	reset portal tcp-cheat statistics	Available in user view

Portal configuration examples

The HP 5500 EI series supports Layer 2 and Layer 3 portal authentication. The HP 5500 SI series supports only Layer 2 portal authentication. Therefore, only the example "[Configuring Layer 2 portal authentication](#)" is applicable to the HP 5500 SI series.

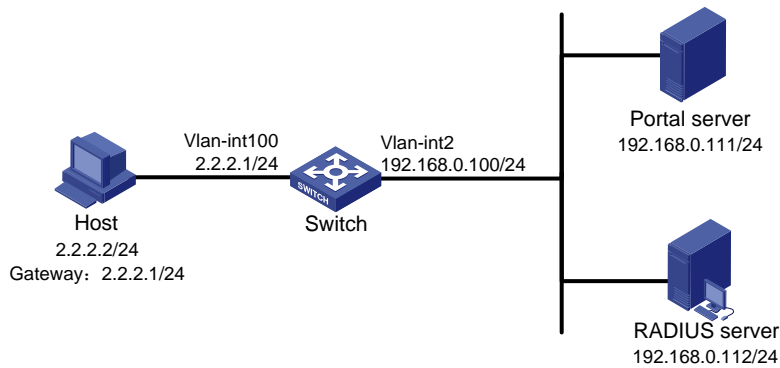
Configuring direct portal authentication

Network requirements

As shown in [Figure 60](#):

- The host is directly connected to the switch and the switch is configured for direct authentication. The host is assigned with a public network IP address either manually or through DHCP. Before passing portal authentication, users can access only the portal server. After passing portal authentication, users can access Internet resources.
- A RADIUS server serves as the authentication, authorization, and accounting server.

Figure 60 Network diagram



Configure IP addresses for the host, switch, and servers as shown in Figure 60 and make sure that they can reach each other.

Configure the RADIUS server properly to provide authentication and accounting functions for users.

Configuring the portal server (IMC PLAT 5.0)

This example assumes that the portal server runs on IMC PLAT 5.0(E0101) and IMC UAM 5.0(E0101).

Configure the portal server.

Log in to IMC and select the **Service** tab. Then, select **User Access Manager > Portal Service Management > Server** from the navigation tree to enter the portal server configuration page, as shown in Figure 61.

- Configure the portal server parameters as needed. This example uses the default settings.

Figure 61 Portal server configuration

The screenshot shows the IMC configuration interface for a Portal Server. The breadcrumb path is Service >> User Access Manager >> Portal Service Management >> Server. The configuration is divided into Basic Information and Advanced Information sections.

Basic Information:

- Log Level: Info
- Request Timeout: 5 Seconds
- Server Heartbeat Interval: 20 Seconds
- User Heartbeat Interval: 5 Minutes
- Portal Page: http://192.168.0.111:8080/portal

Advanced Information:

Service Type List

Add

Total Items: 0.

Service Type ID	Service Type	Delete
-----------------	--------------	--------

OK

Configure the IP address group.

Select **User Access Manager > Portal Service Management > IP Group** from the navigation tree to enter the portal IP address group configuration page. Then, click **Add** to enter the page shown in [Figure 62](#).

- Enter the IP group name.
- Enter the start IP address and end IP address of the IP group. Make sure that the host IP address is in the IP group.
- Select a service group. By default, the group **Ungrouped** is used.
- Select the IP group type **Normal**.

Figure 62 Adding an IP address group

Service >> User Access Manager >> Portal Service Management >> Portal IP Group Configuration >> Add IP Group

Add IP Group

* IP Group Name

* Start IP

* End IP

Service Group ▼

* Action ▼

OK Cancel

Add a portal device.

Select **User Access Manager > Portal Service Management > Device** from the navigation tree to enter the portal device configuration page. Then, click **Add** to enter the page shown in [Figure 63](#).

- Enter the device name **NAS**.
- Enter the IP address of the switch's interface connected to the user.
- Enter the key, which must be the same as that configured on the switch.
- Set whether to enable IP address reallocation. This example uses direct portal authentication, and therefore select **No** from the **Reallocate IP** list.
- Select whether to support sever heartbeat and user heartbeat functions. In this example, select **No** for both **Support Server Heartbeat** and **Support User Heartbeat**.

Figure 63 Adding a portal device

Service >> User Access Manager >> Portal Service Management >> Portal Device Configuration >> Add Device





Add Device

* Device Name	<input type="text" value="NAS"/>	* IP Address	<input type="text" value="2.2.2.1"/>
* Version	<input type="text" value="Portal 2.0"/> ▾	* Key	<input type="text" value="portal"/>
* Listening Port	<input type="text" value="2000"/>	* Local Challenge	<input type="text" value="No"/> ▾
* Authentication Retries	<input type="text" value="2"/>	* Logout Retries	<input type="text" value="4"/>
* Reallocate IP	<input type="text" value="No"/> ▾	* Support User Heartbeat	<input type="text" value="No"/> ▾
* Support Server Heartbeat	<input type="text" value="No"/> ▾		
* Service Group	<input type="text" value="Ungrouped"/> ▾		
Device Description	<input type="text"/>		

Associate the portal device with the IP address group.

As shown in [Figure 64](#), click the icon in the **Port Group Information Management** column of device **NAS** to enter the port group configuration page.

Figure 64 Device list

Device Information List							
<input type="button" value="Add"/>							
1-2 of 2. Page 1 of 1.						Items per Page: 8 15 [50] 100 200	
Device Name	Version	Service Group	IP Address	Port Group Information Management	Details	Modify	Delete
NAS	Portal 2.0	Ungrouped	2.2.2.1				

On the port group configuration page, click **Add** to enter the page shown in [Figure 65](#). Perform the following configurations:

- Enter the port group name.
- Select the configured IP address group. The IP address used by the user to access the network must be within this IP address group.
- Use the default settings for other parameters.

Figure 65 Adding a port group

The screenshot shows the 'Add Port Group Info' configuration window. The fields are as follows:

- Port Group Name: group
- Start Port: 0
- End Port: zzzzzz
- Protocol: HTTP
- NAT or Not: No
- Authentication Type: CHAP
- Heartbeat Interval: 10 Minutes
- Language: Dynamic Detection
- Quick Authentication: No
- Error Transparent Transmission: Yes
- IP Group: Portal_user
- Heartbeat Timeout: 30 Minutes
- User Domain: (empty)
- Port Group Description: (empty)
- User Attribute Type: (empty)
- Default Authentication Type: Web Identity AuthN
- Default Authentication Page: index_default.jsp

Buttons: OK, Cancel

Select **User Access Manager > Service Parameters > Validate System Configuration** from the navigation tree to validate the configurations.

Configuring the switch

1. Configure a RADIUS scheme:

Create a RADIUS scheme named **rs1** and enter its view.

```
<Switch> system-view
[Switch] radius scheme rs1
```

Set the server type for the RADIUS scheme. When using the IMC server, set the server type to **extended**.

```
[Switch-radius-rs1] server-type extended
```

Specify the primary authentication server and primary accounting server, and configure the keys for communication with the servers.

```
[Switch-radius-rs1] primary authentication 192.168.0.112
[Switch-radius-rs1] primary accounting 192.168.0.112
[Switch-radius-rs1] key authentication radius
[Switch-radius-rs1] key accounting radius
```

Specify that the ISP domain name should not be included in the username sent to the RADIUS server.

```
[Switch-radius-rs1] user-name-format without-domain
[Switch-radius-rs1] quit
```

2. Configure an authentication domain:

Create an ISP domain named **dm1** and enter its view.

```
[Switch] domain dm1
```

Configure AAA methods for the ISP domain.

```
[Switch-isp-dm1] authentication portal radius-scheme rs1
[Switch-isp-dm1] authorization portal radius-scheme rs1
[Switch-isp-dm1] accounting portal radius-scheme rs1
[Switch-isp-dm1] quit
```

Configure dm1 as the default ISP domain for all users. Then, if a user enters the username without the ISP domain at logon, the authentication and accounting methods of the default domain are used for the user.

```
[Switch] domain default enable dm1
```

3. Configure portal authentication:

Configure a portal server on the switch, making sure that the IP address, port number and URL match those of the actual portal server.

```
[Switch] portal server newpt ip 192.168.0.111 key portal port 50100 url  
http://192.168.0.111:8080/portal
```

Enable portal authentication on the interface connecting the host.

```
[Switch] interface vlan-interface 100  
[Switch-Vlan-interface100] portal server newpt method direct  
[Switch-Vlan-interface100] quit
```

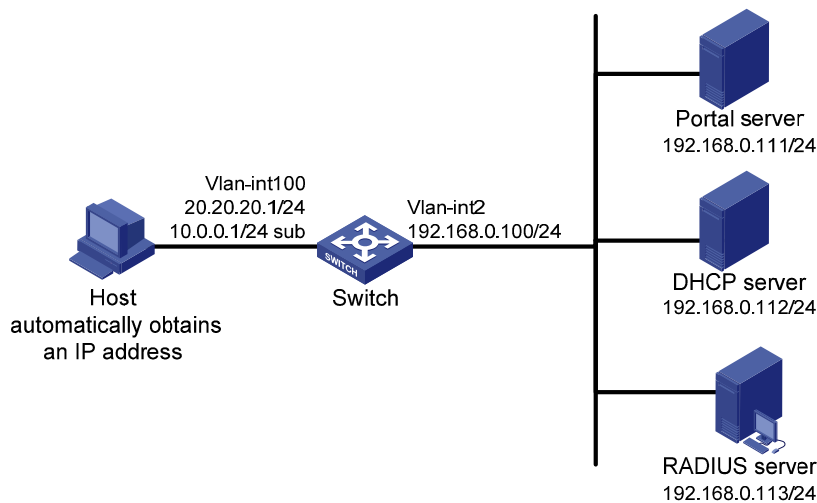
Configuring re-DHCP portal authentication

Network requirements

As shown in Figure 66:

- The host is directly connected to the switch and the switch is configured for re-DHCP authentication. The host is assigned with an IP address through the DHCP server. Before passing portal authentication, the host uses an assigned private IP address. After passing portal authentication, the host can get a public IP address and access Internet resources.
- A RADIUS server serves as the authentication/accounting server.

Figure 66 Network diagram



Configuration procedure

When you configure re-DHCP portal authentication, follow these guidelines:

- Configure a public address pool (20.20.20.0/24, in this example) and a private address pool (10.0.0.0/24, in this example) on the DHCP server. (Details not shown)
- The switch must be configured as a DHCP relay agent and the portal-enabled interface must be configured with a primary IP address (a public IP address) and a secondary IP address (a private

IP address). For information about DHCP relay agent configuration, see *Layer 3—IP Services Configuration Guide*.

- Make sure the IP address of the portal device added on the portal server is the public IP address of the interface connecting users (20.20.20.1 in this example), the private IP address range for the IP address group associated with the portal device is the private network segment where the users reside (10.0.0.0/24 in this example), and the public IP address range for the IP address group is the public network segment 20.20.20.0/24.
- Configure IP addresses for the switch and servers as shown in [Figure 66](#) and make sure that the host, switch, and servers can reach each other.
- Configure the RADIUS server properly to provide authentication and accounting functions for users.

Perform the following configuration to configure re-DHCP authentication on the switch:

1. Configure a RADIUS scheme:

Create a RADIUS scheme named **rs1** and enter its view.

```
<Switch> system-view
```

```
[Switch] radius scheme rs1
```

Set the server type for the RADIUS scheme. When using the IMC server, set the server type to **extended**.

```
[Switch-radius-rs1] server-type extended
```

Specify the primary authentication server and primary accounting server, and configure the keys for communication with the servers.

```
[Switch-radius-rs1] primary authentication 192.168.0.113
```

```
[Switch-radius-rs1] primary accounting 192.168.0.113
```

```
[Switch-radius-rs1] key authentication radius
```

```
[Switch-radius-rs1] key accounting radius
```

Specify that the ISP domain name should not be included in the username sent to the RADIUS server.

```
[Switch-radius-rs1] user-name-format without-domain
```

```
[Switch-radius-rs1] quit
```

2. Configure an authentication domain:

Create an ISP domain named **dm1** and enter its view.

```
[Switch] domain dm1
```

Configure AAA methods for the ISP domain.

```
[Switch-isp-dm1] authentication portal radius-scheme rs1
```

```
[Switch-isp-dm1] authorization portal radius-scheme rs1
```

```
[Switch-isp-dm1] accounting portal radius-scheme rs1
```

```
[Switch-isp-dm1] quit
```

Configure domain **dm1** as the default ISP domain for all users. Then, if a user enters the username without the ISP domain at logon, the authentication and accounting methods of the default domain are used for the user.

```
[Switch] domain default enable dm1
```

3. Configure portal authentication:

Configure the portal server as follows:

- Name: newpt
- IP address: 192.168.0.111
- Key: portal

- Port number: 50100
- URL: `http://192.168.0.111:8080/portal`.

```
[Switch] portal server newpt ip 192.168.0.111 key portal port 50100 url
http://192.168.0.111:8080/portal
```

Configure the switch as a DHCP relay agent, and enable the IP address check function.

```
[Switch] dhcp enable
[Switch] dhcp relay server-group 0 ip 192.168.0.112
[Switch] interface vlan-interface 100
[Switch-Vlan-interface100] ip address 20.20.20.1 255.255.255.0
[Switch-Vlan-interface100] ip address 10.0.0.1 255.255.255.0 sub
[Switch-Vlan-interface100] dhcp select relay
[Switch-Vlan-interface100] dhcp relay server-select 0
[Switch-Vlan-interface100] dhcp relay address-check enable
```

Enable re-DHCP portal authentication on the interface connecting the host.

```
[Switch-Vlan-interface100] portal server newpt method redhcp
[Switch-Vlan-interface100] quit
```

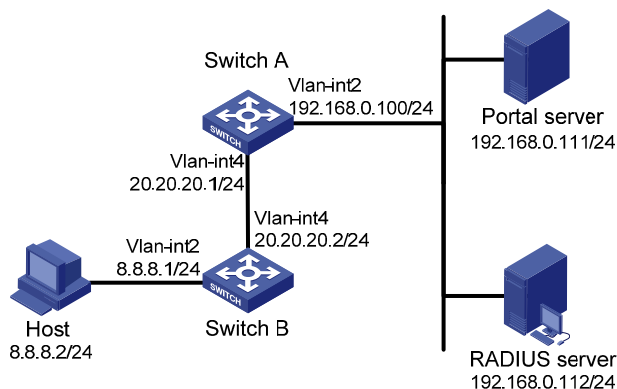
Configuring cross-subnet portal authentication

Network requirements

As shown in [Figure 67](#):

- Switch A is configured for cross-subnet portal authentication. Before passing portal authentication, the host can access only the portal server. After passing portal authentication, the host can access Internet resources.
- The host accesses Switch A through Switch B.
- A RADIUS server serves as the authentication/accounting server.

Figure 67 Network diagram



Configuration procedure

When configuring cross-subnet portal authentication, follow these guidelines:

- Configure IP addresses for the host, switches, and servers as shown in [Figure 67](#) and make sure they can reach each other.
- Configure the RADIUS server properly to provide authentication and accounting functions for users.

- Make sure the IP address of the portal device added on the portal server is the IP address of the interface connecting users (20.20.20.1 in this example), and the IP address group associated with the portal device is the network segment where the users reside (8.8.8.0/24 in this example).

Perform the following configuration to configure cross-subnet portal authentication on Switch A:

1. Configure a RADIUS scheme:

Create a RADIUS scheme named **rs1** and enter its view.

```
<SwitchA> system-view
```

```
[SwitchA] radius scheme rs1
```

Set the server type for the RADIUS scheme. When using the IMC server, set it to **extended**.

```
[SwitchA-radius-rs1] server-type extended
```

Specify the primary authentication server and primary accounting server, and configure the keys for communication with the servers.

```
[SwitchA-radius-rs1] primary authentication 192.168.0.112
```

```
[SwitchA-radius-rs1] primary accounting 192.168.0.112
```

```
[SwitchA-radius-rs1] key authentication radius
```

```
[SwitchA-radius-rs1] key accounting radius
```

Specify that the ISP domain name should not be included in the username sent to the RADIUS server.

```
[SwitchA-radius-rs1] user-name-format without-domain
```

```
[SwitchA-radius-rs1] quit
```

2. Configure an authentication domain:

Create an ISP domain named **dm1** and enter its view.

```
[SwitchA] domain dm1
```

Configure AAA methods for the ISP domain.

```
[SwitchA-isp-dm1] authentication portal radius-scheme rs1
```

```
[SwitchA-isp-dm1] authorization portal radius-scheme rs1
```

```
[SwitchA-isp-dm1] accounting portal radius-scheme rs1
```

```
[SwitchA-isp-dm1] quit
```

Configure domain **dm1** as the default ISP domain for all users. Then, if a user enters the username without the ISP domain at logon, the authentication and accounting methods of the default domain are used for the user.

```
[SwitchA] domain default enable dm1
```

3. Configure portal authentication:

Configure the portal server as follows:

- Name: newpt
- IP address: 192.168.0.111
- Key: portal
- Port number: 50100
- URL: http://192.168.0.111:8080/portal.

```
[SwitchA] portal server newpt ip 192.168.0.111 key portal port 50100 url
```

```
http://192.168.0.111:8080/portal
```

Enable portal authentication on the interface connecting Switch B.

```
[SwitchA] interface vlan-interface 4
```

```
[SwitchA-Vlan-interface4] portal server newpt method layer3
```

```
[SwitchA-Vlan-interface4] quit
```

On Switch B, configure a default route to subnet 192.168.0.0/24, setting the next hop as 20.20.20.1. (Details not shown.)

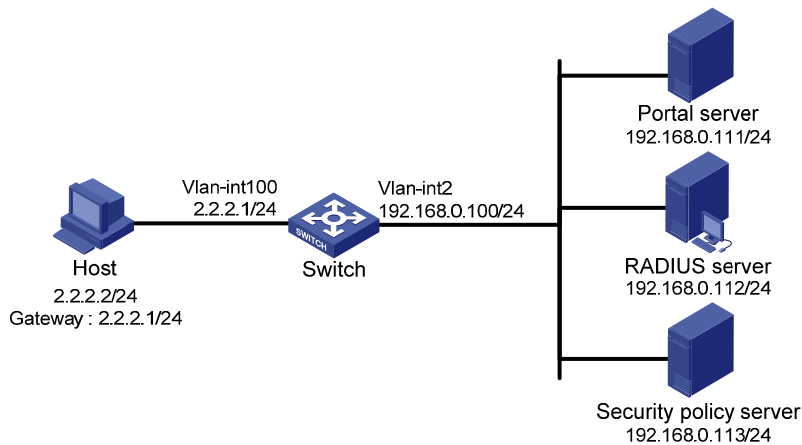
Configuring direct portal authentication with extended functions

Network requirements

As shown in [Figure 68](#):

- The host is directly connected to the switch and the switch is configured for direct extended portal authentication. The host is assigned with a public network IP address either manually or through DHCP. If the host fails security check after passing identity authentication, the host can access only subnet 192.168.0.0/24. After passing security check, the host can access Internet resources.
- A RADIUS server serves as the authentication/accounting server.

Figure 68 Network diagram



Configuration procedure

Configure IP addresses for the host, switch, and servers as shown in [Figure 68](#) and make sure they can reach each other.

Configure the RADIUS server properly to provide authentication and accounting functions for users.

Configure the switch:

1. Configure a RADIUS scheme:

Create a RADIUS scheme named **rs1** and enter its view.

```
<Switch> system-view
[Switch] radius scheme rs1
```

Set the server type for the RADIUS scheme. When using the IMC server, set the server type to **extended**.

```
[Switch-radius-rs1] server-type extended
```

Specify the primary authentication server and primary accounting server, and configure the keys for communication with the servers.

```
[Switch-radius-rs1] primary authentication 192.168.0.112
[Switch-radius-rs1] primary accounting 192.168.0.112
[Switch-radius-rs1] key accounting radius
```

```
[Switch-radius-rs1] key authentication radius
[Switch-radius-rs1] user-name-format without-domain
# Configure the IP address of the security policy server.
[Switch-radius-rs1] security-policy-server 192.168.0.113
[Switch-radius-rs1] quit
```

2. Configure an authentication domain:

Create an ISP domain named **dm1** and enter its view.

```
[Switch] domain dm1
```

Configure AAA methods for the ISP domain.

```
[Switch-isp-dm1] authentication portal radius-scheme rs1
[Switch-isp-dm1] authorization portal radius-scheme rs1
[Switch-isp-dm1] accounting portal radius-scheme rs1
[Switch-isp-dm1] quit
```

Configure domain **dm1** as the default ISP domain for all users. Then, if a user enters the username without the ISP domain at logon, the authentication and accounting methods of the default domain are used for the user.

```
[Switch] domain default enable dm1
```

3. Configure the ACL (ACL 3000) for resources on subnet 192.168.0.0/24 and the ACL (ACL 3001) for Internet resources:

```
[Switch] acl number 3000
[Switch-acl-adv-3000] rule permit ip destination 192.168.0.0 0.0.0.255
[Switch-acl-adv-3000] rule deny ip
[Switch-acl-adv-3000] quit
[Switch] acl number 3001
[Switch-acl-adv-3001] rule permit ip
[Switch-acl-adv-3001] quit
```

On the security policy server, specify ACL 3000 as the isolation ACL and ACL 3001 as the security ACL.

4. Configure portal authentication:

Configure the portal server as follows:

- Name: newpt
- IP address: 192.168.0.111
- Key: portal
- Port number: 50100
- URL: <http://192.168.0.111:8080/portal>.

```
[Switch] portal server newpt ip 192.168.0.111 key portal port 50100 url
http://192.168.0.111:8080/portal
```

Configure a portal-free rule on the interface connecting the portal server.

```
[Switch] portal free-rule 1 source interface ethernet 1/0/1 destination any
[Switch] quit
```

Enable portal authentication on the interface connecting the host.

```
[Switch] interface vlan-interface 100
[Switch-Vlan-interface100] portal server newpt method direct
[Switch-Vlan-interface100] quit
```

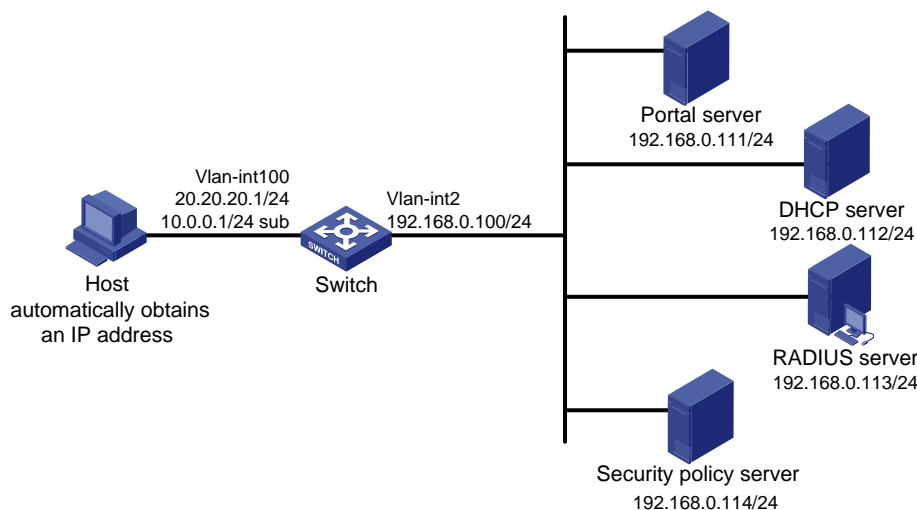
Configuring re-DHCP portal authentication with extended functions

Network requirements

As shown in [Figure 69](#):

- The host is directly connected to the switch and the switch is configured for re-DHCP authentication. The host is assigned with an IP address through the DHCP server. Before passing portal authentication, the host uses an assigned private IP address. After passing portal authentication, the host can get a public IP address.
- If the host fails security check after passing identity authentication, the host can access only subnet 192.168.0.0/24. After passing the security check, the host can access Internet resources.
- A RADIUS server serves as the authentication/accounting server.

Figure 69 Network diagram



Configuration procedure

When you configure re-DHCP portal authentication, follow these guidelines:

- For re-DHCP portal authentication, configure a public address pool (20.20.20.0/24, in this example) and a private address pool (10.0.0.0/24, in this example) on the DHCP server. (Details not shown)
- For re-DHCP portal authentication, the switch must be configured as a DHCP relay agent and the portal-enabled interface must be configured with a primary IP address (a public IP address) and a secondary IP address (a private IP address). For information about DHCP relay agent configuration, see *Layer 3—IP Services Configuration Guide*.
- Make sure the IP address of the portal device added on the portal server is the public IP address of the interface connecting users (20.20.20.1 in this example), the private IP address range for the IP address group associated with the portal device is the private network segment where the users reside (10.0.0.0/24 in this example), and the public IP address range for the IP address group is the public network segment 20.20.20.0/24.
- Configure IP addresses for the switch and servers as shown in [Figure 69](#) and make sure that the host, switch, and servers can reach each other.
- Configure the RADIUS server properly to provide authentication and accounting functions for users.

Perform the following configuration to configure re-DHCP portal authentication with extended functions on the switch:

1. Configure a RADIUS scheme:

Create a RADIUS scheme named **rs1** and enter its view.

```
<Switch> system-view
```

```
[Switch] radius scheme rs1
```

Set the server type for the RADIUS scheme. When using the IMC server, set the server type to **extended**.

```
[Switch-radius-rs1] server-type extended
```

Specify the primary authentication server and primary accounting server, and configure the keys for communication with the servers.

```
[Switch-radius-rs1] primary authentication 192.168.0.113
```

```
[Switch-radius-rs1] primary accounting 192.168.0.113
```

```
[Switch-radius-rs1] key accounting radius
```

```
[Switch-radius-rs1] key authentication radius
```

```
[Switch-radius-rs1] user-name-format without-domain
```

Configure the IP address of the security policy server.

```
[Switch-radius-rs1] security-policy-server 192.168.0.114
```

```
[Switch-radius-rs1] quit
```

2. Configure an authentication domain:

Create an ISP domain named **dm1** and enter its view.

```
[Switch] domain dm1
```

Configure AAA methods for the ISP domain.

```
[Switch-isp-dm1] authentication portal radius-scheme rs1
```

```
[Switch-isp-dm1] authorization portal radius-scheme rs1
```

```
[Switch-isp-dm1] accounting portal radius-scheme rs1
```

```
[Switch-isp-dm1] quit
```

Configure domain **dm1** as the default ISP domain for all users. Then, if a user enters the username without the ISP domain at logon, the authentication and accounting methods of the default domain are used for the user.

```
[Switch] domain default enable dm1
```

3. Configure the ACL (ACL 3000) for resources on subnet 192.168.0.0/24 and the ACL (ACL 3001) for Internet resources:

```
[Switch] acl number 3000
```

```
[Switch-acl-adv-3000] rule permit ip destination 192.168.0.0 0.0.0.255
```

```
[Switch-acl-adv-3000] rule deny ip
```

```
[Switch-acl-adv-3000] quit
```

```
[Switch] acl number 3001
```

```
[Switch-acl-adv-3001] rule permit ip
```

```
[Switch-acl-adv-3001] quit
```

On the security policy server, specify ACL 3000 as the isolation ACL and ACL 3001 as the security ACL.

4. Configure portal authentication:

Configure the portal server as follows:

- o Name: newpt

- IP address: 192.168.0.111
- Key: portal
- Port number: 50100
- URL: http://192.168.0.111:8080/portal.

```
[Switch] portal server newpt ip 192.168.0.111 key portal port 50100
url http://192.168.0.111:8080/portal
```

Configure the switch as a DHCP relay agent, and enable the IP address check function.

```
[Switch] dhcp enable
[Switch] dhcp relay server-group 0 ip 192.168.0.112
[Switch] interface vlan-interface 100
[Switch-Vlan-interface100] ip address 20.20.20.1 255.255.255.0
[Switch-Vlan-interface100] ip address 10.0.0.1 255.255.255.0 sub
[Switch-Vlan-interface100] dhcp select relay
[Switch-Vlan-interface100] dhcp relay server-select 0
[Switch-Vlan-interface100] dhcp relay address-check enable
```

Enable re-DHCP portal authentication on the interface connecting the host.

```
[Switch-Vlan-interface100] portal server newpt method redhcp
[Switch-Vlan-interface100] quit
```

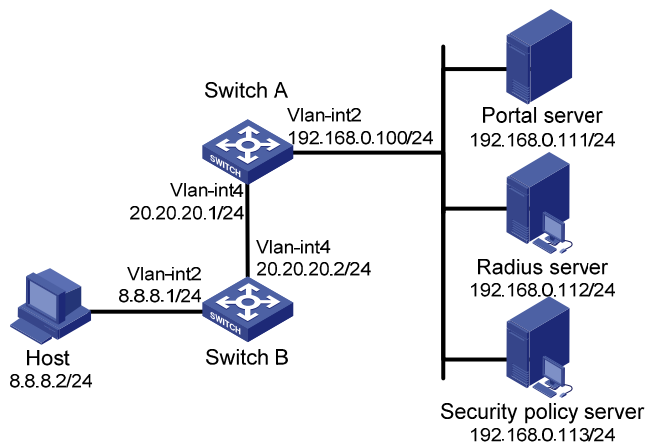
Configuring cross-subnet portal authentication with extended functions

Network requirements

As shown in [Figure 70](#):

- Switch A is configured for cross-subnet extended portal authentication. If the host fails security check after passing identity authentication, the host can access only subnet 192.168.0.0/24. After passing security check, the host can access Internet resources.
- The host accesses Switch A through Switch B.
- A RADIUS server serves as the authentication/accounting server.

Figure 70 Network diagram



Configuration procedure

Make sure the IP address of the portal device added on the portal server is the IP address of the interface connecting users (20.20.20.1 in this example), and the IP address group associated with the portal device is the network segment where the users reside (8.8.8.0/24 in this example).

Configure IP addresses for the host, switches, and servers as shown in [Figure 70](#) and make sure that they can reach each other.

Configure the RADIUS server properly to provide authentication and accounting functions for users.

Configure Switch A:

1. Configure a RADIUS scheme:

Create a RADIUS scheme named **rs1** and enter its view.

```
<SwitchA> system-view
```

```
[SwitchA] radius scheme rs1
```

Set the server type for the RADIUS scheme. When using the IMC server, set the server type to **extended**.

```
[SwitchA-radius-rs1] server-type extended
```

Specify the primary authentication server and primary accounting server, and configure the keys for communication with the servers.

```
[SwitchA-radius-rs1] primary authentication 192.168.0.112
```

```
[SwitchA-radius-rs1] primary accounting 192.168.0.112
```

```
[SwitchA-radius-rs1] key accounting radius
```

```
[SwitchA-radius-rs1] key authentication radius
```

```
[SwitchA-radius-rs1] user-name-format without-domain
```

Configure the IP address of the security policy server.

```
[SwitchA-radius-rs1] security-policy-server 192.168.0.113
```

```
[SwitchA-radius-rs1] quit
```

2. Configure an authentication domain:

Create an ISP domain named **dm1** and enter its view.

```
[SwitchA] domain dm1
```

Configure AAA methods for the ISP domain.

```
[SwitchA-isp-dm1] authentication portal radius-scheme rs1
```

```
[SwitchA-isp-dm1] authorization portal radius-scheme rs1
```

```
[SwitchA-isp-dm1] accounting portal radius-scheme rs1
```

```
[SwitchA-isp-dm1] quit
```

Configure domain **dm1** as the default ISP domain for all users. Then, if a user enters the username without the ISP domain at logon, the authentication and accounting methods of the default domain are used for the user.

```
[SwitchA] domain default enable dm1
```

3. Configure the ACL (ACL 3000) for resources on subnet 192.168.0.0/24 and the ACL (ACL 3001) for Internet resources:

```
[SwitchA] acl number 3000
```

```
[SwitchA-acl-adv-3000] rule permit ip destination 192.168.0.0 0.0.0.255
```

```
[SwitchA-acl-adv-3000] rule deny ip
```

```
[SwitchA-acl-adv-3000] quit
```

```
[SwitchA] acl number 3001
```

```
[SwitchA-acl-adv-3001] rule permit ip
```



```
[SwitchA-acl-adv-3001] quit
```

On the security policy server, specify ACL 3000 as the isolation ACL and ACL 3001 as the security ACL.

4. Configure portal authentication

Configure the portal server as follows:

- o Name: newpt
- o IP address: 192.168.0.111
- o Key: portal
- o Port number: 50100
- o URL: http://192.168.0.111:8080/portal.

```
[SwitchA] portal server newpt ip 192.168.0.111 key portal port 50100 url  
http://192.168.0.111:8080/portal
```

Enable portal authentication on the interface connecting Switch B.

```
[SwitchA] interface vlan-interface 4  
[SwitchA-Vlan-interface4] portal server newpt method layer3  
[SwitchA-Vlan-interface4] quit
```

On Switch B, configure a default route to subnet 192.168.0.0/24, setting the next hop as 20.20.20.1. (Details not shown.)

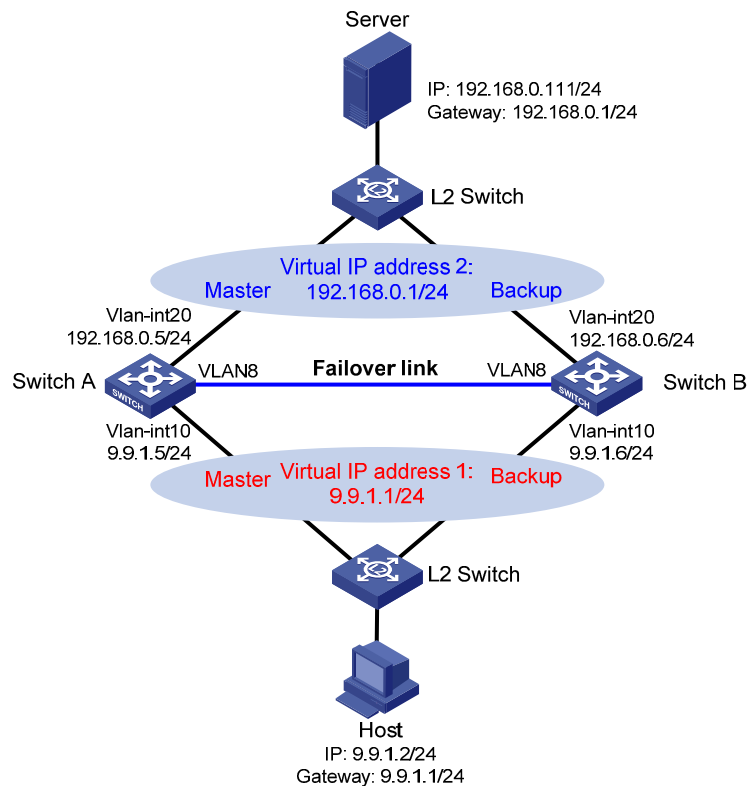
Configuring portal stateful failover

Network requirements

As shown in [Figure 71](#), a failover link is present between Switch A and Switch B. Both Switch A and Switch B support portal authentication. Configure stateful failover between Switch A and Switch B to support portal service backup and use VRRP to implement traffic switchover between the switches. More specifically,

- When Switch A works normally, Host accesses Switch A for portal authentication before accessing the Internet; when Switch A fails, Host accesses the Internet through Switch B. The VRRP uplink/downlink detection mechanism is used to ensure non-stop traffic forwarding.
- Use the RADIUS server as the authentication/accounting server. In this example, Server takes the responsibilities of the portal server and the RADIUS server.
- Switch A and Switch B use the failover link to transmit stateful failover related packets and specify VLAN 8 on the switches as the VLAN dedicated for stateful failover related packets.

Figure 71 Network diagram



Configure IP addresses for the host, server, and switches as shown in [Figure 71](#) and make sure that they can reach to each other.

Make sure that Host can access the authentication server through Switch A and Switch B.

Configure VRRP group 1 and VRRP group 2 to implement backup for downstream and upstream links, respectively. For more information about VRRP, see *High Availability Configuration Guide*.

For information about stateful failover configuration, see *High Availability Configuration Guide*.

Configuring the portal server (IMC PLAT 5.0)

This example assumes that the portal server runs on IMC PLAT 5.0(E0101) and IMC UAM 5.0(E0101).

Configure the portal server.

Log in to IMC and select the **Service** tab. Then, select **User Access Manager > Portal Service Management > Server** from the navigation tree to enter the portal server configuration page, as shown in [Figure 72](#).

Configure the portal server parameters as needed. This example uses the default settings.

Figure 72 Portal server configuration

Service >> User Access Manager >> Portal Service Management >> Server

Portal Server Configuration

Basic Information

* Log Level: Info (dropdown) * Request Timeout: 5 (input) Seconds (?)

* Server Heartbeat Interval: 20 (input) Seconds (?) * User Heartbeat Interval: 5 (input) Minutes (?)

Portal Page:

Advanced Information

Service Type List

Total Items: 0.

Service Type ID	Service Type	Delete
OK		

Configure the IP address group.

Select **User Access Manager > Portal Service Management > IP Group** from the navigation tree to enter the portal IP address group configuration page. Then, click **Add** to enter the page shown in [Figure 62](#).

- Enter the IP group name.
- Enter the start IP address and end IP address of the IP group. Make sure that the host IP address is in the IP group.
- Select a service group. By default, the group **Ungrouped** is used.
- Select the IP group type **Normal**.

Figure 73 Adding an IP address group

Service >> User Access Manager >> Portal Service Management >> Portal IP Group Configuration >> Add IP Group

Add IP Group

* IP Group Name:

* Start IP:

* End IP:

Service Group: Ungrouped (dropdown)

* Action: Normal (dropdown)

Add a portal device.

Select **User Access Manager > Portal Service Management > Device** from the navigation tree to enter the portal device configuration page. Then, click **Add** to enter the page shown in [Figure 63](#).

- Enter the device name **NAS**.
- Enter the virtual IP address of the VRRP group that holds the portal-enabled interface.
- Enter the key, which must be the same as that configured on the switch.
- Set whether to enable IP address reallocation. This example uses direct portal authentication, and therefore select **No** from the **Reallocate IP** list.
- Select whether to support sever heartbeat and user heartbeat functions. In this example, select **No** for both **Support Server Heartbeat** and **Support User Heartbeat**.

Figure 74 Adding a portal device

Service >> User Access Manager >> Portal Service Management >> Portal Device Configuration >> Add Device

Add Device

* Device Name	<input type="text" value="NAS"/>	* IP Address	<input type="text" value="9.9.1.1"/>
* Version	<input type="text" value="Portal 2.0"/>	* Key	<input type="text" value="portal"/>
* Listening Port	<input type="text" value="2000"/>	* Local Challenge	<input type="text" value="No"/>
* Authentication Retries	<input type="text" value="2"/>	* Logout Retries	<input type="text" value="4"/>
* Reallocate IP	<input type="text" value="No"/>	* Support Server Heartbeat	<input type="text" value="No"/>
* Support User Heartbeat	<input type="text" value="No"/>	* Support User Heartbeat	<input type="text" value="No"/>
* Service Group	<input type="text" value="Ungrouped"/>		
Device Description	<input type="text"/>		

Associate the portal device with the IP address group.

As shown in [Figure 64](#), click the icon in the **Port Group Information Management** column of device **NAS** to enter the port group configuration page.

Figure 75 Device list

Device Information List							
<input type="button" value="Add"/>							
1-1 of 1. Page 1 of 1.						Items per Page: 3 15 50 100 200	
Device Name	Version	Service Group	IP Address	Port Group Information Management	Details	Modify	Delete
NAS	Portal 2.0	Ungrouped	9.9.1.1				

On the port group configuration page, click **Add** to enter the page shown in [Figure 65](#). Perform the following configurations:

- Enter the port group name.
- Select the configured IP address group. The IP address used by the user to access the network must be within this IP address group.
- Use the default settings for other parameters.

Figure 76 Adding a port group

Service >> User Access Manager >> Portal Service Management >> Portal Device Configuration >> Port Group Info Config >> Add ? Help

Port Group Info

Add Port Group Info			
* Port Group Name	group	* Language	Dynamic Detection
* Start Port	0	* End Port	zzzzz
* Protocol	HTTP	* Quick Authentication	No
* NAT or Not	No	* Error Transparent Transmission	Yes
* Authentication Type	CHAP	* IP Group	Portal_user
* Heartbeat Interval	10	* Heartbeat Timeout	30
User Domain		Port Group Description	
User Attribute Type		Default Authentication Page	index_default.jsp
Default Authentication Type	Web Identity AuthN		

OK Cancel

Select **User Access Manager > Service Parameters > Validate System Configuration** from the navigation tree to validate the configurations.

Configuring Switch A

1. Configure VRRP:

Create VRRP group 1, and configure the virtual IP address of the VRRP group 1 as 9.9.1.1.

```
<SwitchA> system-view
[SwitchA] interface vlan-interface 10
[SwitchA-Vlan-interface10] vrrp vrid 1 virtual-ip 9.9.1.1
```

Set the priority of VLAN-interface 10 in VRRP group 1 to 200.

```
[SwitchA-Vlan-interface10] vrrp vrid 1 priority 200
```

On VLAN-interface 10, configure the interface to be tracked as VLAN-interface 20 and reduce the priority of VLAN-interface 10 in VRRP group 1 by 150 when the interface state of VLAN-interface 20 becomes Down or Removed.

```
[SwitchA-Vlan-interface10] vrrp vrid 1 track interface vlan-interface20 reduced 150
[SwitchA-Vlan-interface10] quit
```

Create VRRP group 2, and configure the virtual IP address of the VRRP group 2 as 192.168.0.1.

```
[SwitchA] interface vlan-interface 20
[SwitchA-Vlan-interface20] vrrp vrid 2 virtual-ip 192.168.0.1
```

Set the priority of VLAN-interface 20 in VRRP group 2 to 200.

```
[SwitchA-Vlan-interface20] vrrp vrid 2 priority 200
```

On VLAN-interface 20, configure the interface to be tracked as VLAN-interface 10 and reduce the priority of VLAN-interface 20 in VRRP group 2 by 150 when the interface state of VLAN-interface 10 becomes Down or Removed.

```
[SwitchA-Vlan-interface20] vrrp vrid 2 track interface vlan-interface10 reduced 150
[SwitchA-Vlan-interface20] quit
```

2. Configure a RADIUS scheme:

Create RADIUS scheme **rs1** and enter its view.

```
[SwitchA] radius scheme rs1
```

Configure the server type for the RADIUS scheme. When using the IMC server, configure the RADIUS server type as **extended**.

```
[SwitchA-radius-rs1] server-type extended
```

Specify the primary authentication server and primary accounting server, and configure the keys for communication with the servers.

```
[SwitchA-radius-rs1] primary authentication 192.168.0.111
```

```
[SwitchA-radius-rs1] primary accounting 192.168.0.111
```

```
[SwitchA-radius-rs1] key authentication expert
```

```
[SwitchA-radius-rs1] key accounting expert
```

Configure the access device to not carry the ISP domain name in the username sent to the RADIUS server. (Optional, configure the username format as needed.)

```
[SwitchA-radius-rs1] user-name-format without-domain
```

```
[SwitchA-radius-rs1] quit
```

3. Configure an authentication domain:

Create ISP domain **dm1** and enter its view.

```
[SwitchA] domain dm1
```

Configure AAA methods for the ISP domain.

```
[SwitchA-isp-dm1] authentication portal radius-scheme rs1
```

```
[SwitchA-isp-dm1] authorization portal radius-scheme rs1
```

```
[SwitchA-isp-dm1] accounting portal radius-scheme rs1
```

```
[SwitchA-isp-dm1] quit
```

Configure domain **dm1** as the default ISP domain for all users. Then, if a user enters a username without any ISP domain at logon, the authentication and accounting methods of the default domain are used for the user.

```
[SwitchA] domain default enable dm1
```

4. Enable portal authentication on the interface connecting the host:

Configure a portal server on the switch, making sure that the IP address, port number and URL match those of the actual portal server.

```
[SwitchA] portal server newpt ip 192.168.0.111 key portal port 50100 url
```

```
http://192.168.0.111:8080/portal
```

Enable portal authentication on the interface connecting the host.

```
[SwitchA] interface vlan-interface 10
```

```
[SwitchA-Vlan-interface10] portal server newpt method layer3
```

Specify the source IP address of outgoing portal packets as 9.9.1.1, the virtual IP address of VRRP group 1.

```
[SwitchA-Vlan-interface10] portal nas-ip 9.9.1.1
```

5. Configure portal stateful failover:

Assign interface VLAN-interface 10 to portal group 1.

```
[SwitchA-Vlan-interface10] portal backup-group 1
```

```
[SwitchA-Vlan-interface10] quit
```

Set the device ID for Switch A in stateful failover mode to 1.

```
[SwitchA] nas device-id 1
```

Specify the source IP address of outgoing RADIUS packets as 192.168.0.1, the virtual IP address of VRRP group 2.

```
[SwitchA] radius nas-ip 192.168.0.1
```

Make sure you have added the access device with IP address 192.168.0.1 on the RADIUS server.

6. Configure the stateful failover function:

```
# Configure the VLAN for stateful failover as VLAN 8.
[SwitchA] dnbk vlan 8
# Enable stateful failover and configure it to support the symmetric path.
[SwitchA] dnbk enable backup-type symmetric-path
```

Configuring Switch B

1. Configure VRRP:

```
# Create VRRP group 1, and configure the virtual IP address of the VRRP group 1 as 9.9.1.1.
<SwitchB> system-view
[SwitchB] interface vlan-interface 10
[SwitchB-Vlan-interface10] vrrp vrid 1 virtual-ip 9.9.1.1
# Set the priority of VLAN-interface 10 in VRRP group 1 to 150.
[SwitchB-Vlan-interface10] vrrp vrid 1 priority 150
[SwitchB-Vlan-interface10] quit
# Create VRRP group 2, and configure the virtual IP address of the VRRP group 2 as 192.168.0.1.
[SwitchB] interface vlan-interface 20
[SwitchB-Vlan-interface20] vrrp vrid 2 virtual-ip 192.168.0.1
# Set the priority of VLAN-interface 20 in VRRP group 2 to 150.
[SwitchB-Vlan-interface20] vrrp vrid 2 priority 150
[SwitchB-Vlan-interface20] quit
```

2. Configure a RADIUS scheme:

```
# Create RADIUS scheme rs1 and enter its view.
[SwitchB] radius scheme rs1
# Configure the server type for the RADIUS scheme. When using the IMC server, configure the
RADIUS server type as extended.
[SwitchB-radius-rs1] server-type extended
# Specify the primary authentication server and primary accounting server, and configure the keys
for communication with the servers.
[SwitchB-radius-rs1] primary authentication 192.168.0.111
[SwitchB-radius-rs1] primary accounting 192.168.0.111
[SwitchB-radius-rs1] key authentication expert
[SwitchB-radius-rs1] key accounting expert
# Configure the access device to not carry the ISP domain name in the username sent to the
RADIUS server. (Optional, configure the username format as needed.)
[SwitchB-radius-rs1] user-name-format without-domain
[SwitchB-radius-rs1] quit
```

3. Configure an authentication domain:

```
# Create ISP domain dm1 and enter its view.
[SwitchB] domain dm1
# Configure AAA methods for the ISP domain.
[SwitchB-isp-dm1] authentication portal radius-scheme rs1
[SwitchB-isp-dm1] authorization portal radius-scheme rs1
[SwitchB-isp-dm1] accounting portal radius-scheme rs1
[SwitchB-isp-dm1] quit
```

Configure domain **dm1** as the default ISP domain for all users. Then, if a user enters a username without any ISP domain at logon, the authentication and accounting methods of the default domain are used for the user.

```
[SwitchB] domain default enable dm1
```

4. Enable portal authentication on the interface connecting the host:

Configure the portal server as needed.

```
[SwitchB] portal server newpt ip 192.168.0.111 key portal port 50100 url
http://192.168.0.111:8080/portal
```

Enable portal authentication on the interface connecting the host.

```
[SwitchB] interface vlan-interface 10
```

```
[SwitchB-Vlan-interface10] portal server newpt method layer3
```

Specify the source IP address of outgoing portal packets as 9.9.1.1, the virtual IP address of VRRP group 1.

```
[SwitchA-Vlan-interface10] portal nas-ip 9.9.1.1
```

5. Configure portal stateful failover:

Assign interface VLAN-interface 10 to portal group 1.

```
[SwitchB-Vlan-interface10] portal backup-group 1
```

```
[SwitchB-Vlan-interface10] quit
```

Set the ID of the device in the stateful failover mode to 2.

```
[SwitchB] nas device-id 2
```

Specify the source IP address of outgoing RADIUS packets as 192.168.0.1, the virtual IP address of VRRP group 2.

```
[SwitchB] radius nas-backup-ip 192.168.0.1
```

Make sure you have added the access device with IP address 192.168.0.1 on the RADIUS server.

6. Configure stateful failure:

Configure the VLAN for stateful failover as VLAN 8.

```
[SwitchB] dmbk vlan 8
```

Enable stateful failover and configure it to support the symmetric path.

```
[SwitchB] dmbk enable backup-type symmetric-path
```

Verifying the configuration

After user Host logs in through Switch A, display the user authentication information by using the **display portal user** command on Switch A and Switch B.

```
[SwitchA] display portal user all
```

```
Index:3
```

```
State:ONLINE
```

```
SubState:NONE
```

```
ACL:NONE
```

```
Work-mode: primary
```

```
VPN instance:NONE
```

```
MAC                IP                Vlan  Interface
```

```
-----
```

```
000d-88f8-0eac      9.9.1.2           10    Vlan-interface10
```

```
Total 1 user(s) matched, 1 listed.
```

```
[SwitchB] display portal user all
```

```
Index:2
```



```

State:ONLINE
SubState:NONE
ACL:NONE
Work-mode: secondary
VPN instance:NONE
MAC                IP                Vlan  Interface
-----
000d-88f8-0eac     9.9.1.2          10    Vlan-interface10
Total 1 user(s) matched, 1 listed.

```

The output shows that the information of user Host is saved on both Switch A and Switch B. The user's working mode on Switch A is primary, and that on Switch B is secondary, which indicate that the user logged in through Switch A and the user information on Switch B was synchronized from Switch A.

Configuring portal server detection and portal user information synchronization

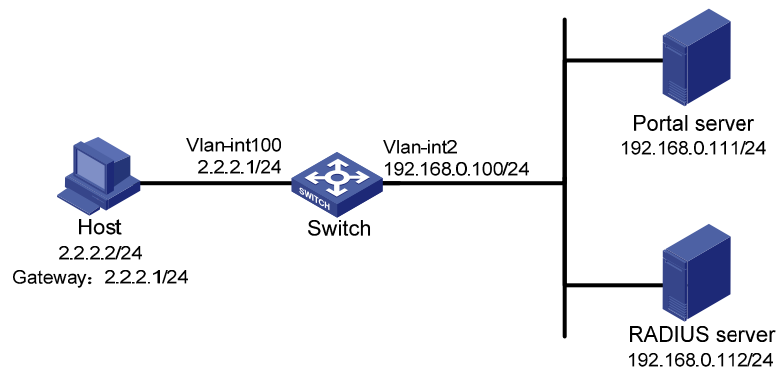
Network requirements

As shown in [Figure 77](#), a host is directly connected to a switch (the access device) and must pass portal authentication before it can access the Internet. A RADIUS server serves as the authentication/accounting server.

Detailed requirements are as follows:

- The host is assigned with a public network IP address either manually or through DHCP. Before passing portal authentication, the host can access only the portal server. After passing portal authentication, the host can access the Internet.
- The access device (Switch) can detect whether the portal server is reachable and send trap messages upon state changes. When the portal server is unreachable due to a connection failure, network device failure, or portal server failure, for example, the access device can disable portal authentication, allowing users to access the Internet without authentication.
- The access device can synchronize portal user information with the portal server periodically.

Figure 77 Network diagram



Configuration considerations

1. Configure the portal server and enable portal server heartbeat function and the portal user heartbeat function.
2. Configure the RADIUS server to implement authentication and accounting.

3. Configure direct portal authentication on interface VLAN-interface 100, which is connected with the user host.
4. Configure the portal server detection function on the access device, so that the access device can detect the status of the portal server by cooperating with the portal server heartbeat function.
5. Configure the portal user information synchronization function, so that the access device can synchronize portal user information with the portal server by cooperating with the portal user heartbeat function.

Configure IP addresses for the host, switch, and servers as shown in [Figure 77](#) and make sure that they can reach each other.

Configure the RADIUS server properly to provide authentication and accounting functions for users.

Configuring the portal server (IMC PLAT 5.0)

This example assumes that the portal server runs IMC PLAT 5.0(E0101) and IMC UAM 5.0(E0101).

Configure the portal server.

Log in to IMC and select the **Service** tab. Then, select **User Access Manager > Portal Service Management > Server** from the navigation tree to enter the portal server configuration page, as shown in [Figure 78](#).

- Configure the portal server heartbeat interval and user heartbeat interval.
- Use the default value for other parameters.

Figure 78 Portal server configuration

Service >> User Access Manager >> Portal Service Management >> Server

Portal Server Configuration

Basic Information

* Log Level: Info

* Request Timeout: 5 Seconds

* Server Heartbeat Interval: 20 Seconds

* User Heartbeat Interval: 5 Minutes

Portal Page: http://192.168.0.111:8080/portal

Advanced Information

Service Type List

Add

Total Items: 0.

Service Type ID	Service Type	Delete
-----------------	--------------	--------

OK

Configure the IP address group.

Select **User Access Manager > Portal Service Management > IP Group** from the navigation tree to enter the portal IP address group configuration page. Then, click **Add** to enter the page shown in [Figure 62](#).

- Enter the IP group name.

- Enter the start IP address and end IP address of the IP group. Make sure that the host IP address is in the IP group.
- Select a service group. By default, the group **Ungrouped** is used.
- Select the IP group type **Normal**.

Figure 79 Adding an IP address group

Service >> User Access Manager >> Portal Service Management >> Portal IP Group Configuration >> Add IP Group

Add IP Group

* IP Group Name	<input type="text" value="Portal_user"/>
* Start IP	<input type="text" value="2.2.2.1"/>
* End IP	<input type="text" value="2.2.2.255"/>
Service Group	<input type="text" value="Ungrouped"/> ▼
* Action	<input type="text" value="Normal"/> ▼

Add a portal device.

Select **User Access Manager > Portal Service Management > Device** from the navigation tree to enter the portal device configuration page. Then, click **Add** to enter the page shown in [Figure 63](#).

- Enter the device name **NAS**.
- Enter the IP address of the switch's interface connected to the user.
- Enter the key, which must be the same as that configured on the switch.
- Set whether to enable IP address reallocation. This example uses direct portal authentication, and therefore select **No** from the **Reallocate IP** list.
- Set whether to support the portal server heartbeat and user heartbeat functions. In this example, select **Yes** for both **Support Server Heartbeat** and **Support User Heartbeat**.

Figure 80 Adding a portal device

Service >> User Access Manager >> Portal Service Management >> Portal Device Configuration >> Add Device





Add Device

* Device Name	<input type="text" value="NAS"/>	* IP Address	<input type="text" value="2.2.2.1"/>
* Version	<input type="text" value="Portal 2.0"/> ▼	* Key	<input type="text" value="portal"/>
* Listening Port	<input type="text" value="2000"/>	* Local Challenge	<input type="text" value="No"/> ▼
* Authentication Retries	<input type="text" value="2"/>	* Logout Retries	<input type="text" value="4"/>
* Reallocate IP	<input type="text" value="No"/> ▼	* Support Server Heartbeat	<input type="text" value="Yes"/> ▼
* Support User Heartbeat	<input type="text" value="Yes"/> ▼	* Support User Heartbeat	<input type="text" value="Yes"/> ▼
* Service Group	<input type="text" value="Ungrouped"/> ▼		
Device Description	<input type="text"/>		

Associate the portal device with the IP address group.

As shown in [Figure 64](#), click the icon in the **Port Group Information Management** column of device **NAS** to enter the port group configuration page.


Figure 81 Device list

Device Information List							
Add							
1-2 of 2. Page 1 of 1.						Items per Page: 8 15 [50] 100 200	
Device Name	Version	Service Group	IP Address	Port Group Information Management	Details	Modify	Delete
NAS	Portal 2.0	Ungrouped	2.2.2.1				

On the port group configuration page, click **Add** to enter the page shown in [Figure 65](#). Perform the following configurations:

- Enter the port group name.
- Select the configured IP address group. The IP address used by the user to access the network must be within this IP address group.
- Use the default settings for other parameters.

Figure 82 Adding a port group

Service >> User Access Manager >> Portal Service Management >> Portal Device Configuration >> Port Group Info Config >> Add 

Port Group Info

Add Port Group Info

* Port Group Name	<input type="text" value="group"/>	* Language	Dynamic Detection <input type="button" value="v"/>
* Start Port	<input type="text" value="0"/>	* End Port	<input type="text" value="zzzzz"/>
* Protocol	HTTP <input type="button" value="v"/>	* Quick Authentication	No <input type="button" value="v"/>
* NAT or Not	No <input type="button" value="v"/>	* Error Transparent Transmission	Yes <input type="button" value="v"/>
* Authentication Type	CHAP <input type="button" value="v"/>	* IP Group	Portal_user <input type="button" value="v"/>
* Heartbeat Interval	<input type="text" value="10"/> Minutes	* Heartbeat Timeout	<input type="text" value="30"/> Minutes
User Domain	<input type="text"/>	Port Group Description	<input type="text"/>
User Attribute Type	<input type="button" value="v"/>	Default Authentication Page	<input type="text" value="index_default.jsp"/>
Default Authentication Type	Web Identity AuthN <input type="button" value="v"/>		

Select **User Access Manager > Service Parameters > Validate System Configuration** from the navigation tree to validate the configurations.

Configure the switch

1. Configure a RADIUS scheme:

Create RADIUS scheme **rs1** and enter its view.

```
<Switch> system-view
[Switch] radius scheme rs1
```

Configure the server type for the RADIUS scheme. When using the IMC server, configure the RADIUS server type as **extended**.

```
[Switch-radius-rs1] server-type extended
```

Specify the primary authentication server and primary accounting server, and configure the keys for communication with the servers.

```
[Switch-radius-rs1] primary authentication 192.168.0.112
[Switch-radius-rs1] primary accounting 192.168.0.112
[Switch-radius-rs1] key authentication radius
[Switch-radius-rs1] key accounting radius
```

Configure the access device to not carry the ISP domain name in the username sent to the RADIUS server.

```
[Switch-radius-rs1] user-name-format without-domain
[Switch-radius-rs1] quit
```

2. Configure an authentication domain:

Create ISP domain **dm1** and enter its view.

```
[Switch] domain dm1
```

Configure AAA methods for the ISP domain.

```
[Switch-isp-dm1] authentication portal radius-scheme rs1
[Switch-isp-dm1] authorization portal radius-scheme rs1
[Switch-isp-dm1] accounting portal radius-scheme rs1
[Switch-isp-dm1] quit
```

Configure domain **dm1** as the default ISP domain for all users. Then, if a user enters a username without the ISP domain at logon, the authentication and accounting methods of the default domain are used for the user.

```
[Switch] domain default enable dm1
```

3. Configure portal authentication:

Configure a portal server on the switch, making sure that the IP address, port number and URL match those of the actual portal server.

```
[Switch] portal server newpt ip 192.168.0.111 key portal port 50100 url
http://192.168.0.111:8080/portal
```

Enable portal authentication on the interface connecting the host.

```
[Switch] interface vlan-interface 100
[Switch-Vlan-interface100] portal server newpt method direct
[Switch-Vlan-interface100] quit
```

4. Configure the portal server detection function:

Configure the access device to detect portal server **newpt**, specifying the detection method as portal heartbeat probe, setting the server probe interval to 40 seconds, and specifying the access device to send a server unreachable trap message and disable portal authentication to permit unauthenticated portal users if two consecutive probes fail.

```
[Switch] portal server newpt server-detect method portal-heartbeat action trap
permit-all interval 40 retry 2
```

The product of **interval** and **retry** must be greater than or equal to the portal server heartbeat interval, and HP recommends configuring the **interval** as a value greater than the portal server heartbeat interval configured on the portal server.

5. Configure portal user synchronization:

Configure the access device to synchronize portal user information with portal server **newpt**, setting the synchronization probe interval to 600 seconds, and specifying the access device to log off users if the users do not appear in the user synchronization packets sent from the server in two consecutive probe intervals.

```
[Switch] portal server newpt user-sync interval 600 retry 2
```

The product of **interval** and **retry** must be greater than or equal to the portal user heartbeat interval, and HP recommends configuring the **interval** as a value greater than the portal user heartbeat interval configured on the portal server.

Verifying the configuration

Use the following command to view information about the portal server:

```
<Switch> display portal server newpt
Portal server:
 1)newpt:
   IP   : 192.168.0.111
   Key  : *****
   Port : 50100
   URL  : http://192.168.0.111:8080/portal
   Status : Up
```

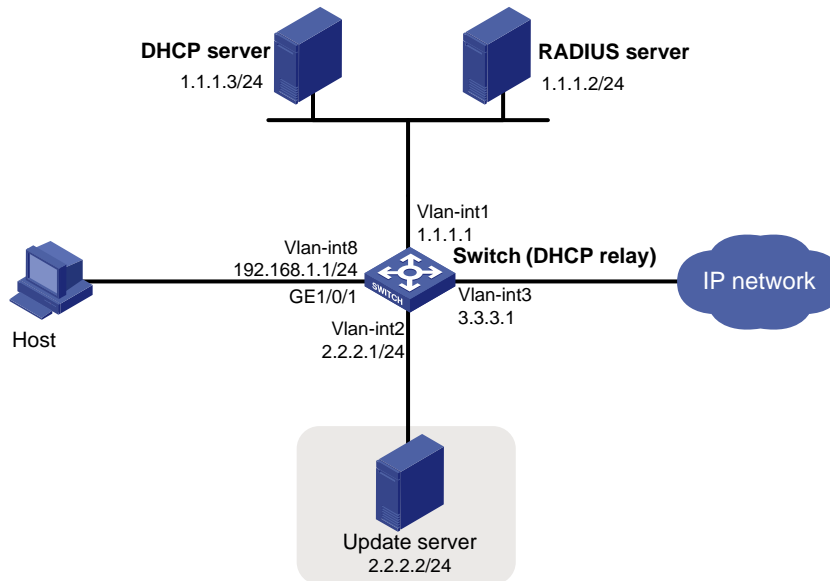
Configuring Layer 2 portal authentication

Network requirements

As shown in [Figure 83](#), a host is directly connected to a switch. The switch performs Layer 2 portal authentication on users connected to port GigabitEthernet 1/0/1. More specifically,

- Use the remote RADIUS server for authentication, authorization and accounting.
- Use the remote DHCP server to assign IP addresses to users.
- The listening IP address of the local portal server is 4.4.4.4. The local portal server pushes the user-defined authentication pages to users and uses HTTPS to transmit authentication data.
- Add users passing authentication to VLAN 3.
- Add users failing authentication to VLAN 2, to allow the users to access resources on the update server.
- The host obtains an IP address through DHCP. Before authentication, the DHCP server assigns an IP address in segment 192.168.1.0/24 to the host. When the host passes the authentication, the DHCP server assigns an IP address in segment 3.3.3.0/24 to the host. When the host fails authentication, the DHCP server assigns an IP address in segment 2.2.2.0/24 to the host.

Figure 83 Network diagram



Configuration procedures

Follow these guidelines to configure Layer 2 portal authentication:

- Make sure that the host, switch, and servers can reach each other before portal authentication is enabled.
- Configure the RADIUS server properly to provide normal authentication/authorization/accounting functions for users. In this example, you must create a portal user account with the account name **userpt** on the RADIUS server, and configure an authorized VLAN for the account.
- On the DHCP server, you must specify the IP address ranges (192.168.1.0/24, 3.3.3.0/24, 2.2.2.0/24), specify the default gateway addresses (192.168.1.1, 3.3.3.1, 2.2.2.1), exclude the update server's address 2.2.2.2 from the address ranges for address allocation, specify the leases for the assigned IP addresses and make sure there is a route to the host. To shorten the IP address update time in case of an authentication state change, set a short lease for each address.
- Because the DHCP server and the DHCP client are not in the same subnet, you need to configure a DHCP relay agent on the subnet of the client. For more information about DHCP relay agent, see *Layer 3—IP Services Configuration Guide*.

Perform the following configuration on the switch to implement Layer 2 portal authentication:

1. Configure portal authentication:

Add Ethernet ports to related VLANs and configure IP addresses for the VLAN interfaces. (Details not shown.)

Configure PKI domain **pkidm**, and apply for a local certificate and CA certificate. For more configuration information, see "[Configuring PKI](#)."

Edit the user-defined authentication pages file, compress it into a zip file named **defaultfile**, and save the file in the root directory of the access device.

Configure SSL server policy **sslsvr**, and specify to use PKI domain **pkidm**.

```
<Switch> system-view
[Switch] ssl server-policy sslsvr
[Switch-ssl-server-policy-sslsvr] pki pkidm
[Switch-ssl-server-policy-sslsvr] quit
```

```

# Configure the local portal server to support HTTPS and reference SSL server policy sslsvr.
[Switch] portal local-server https server-policy sslsvr
# Configure the IP address of loopback interface 12 as 4.4.4.4.
[Switch] interface loopback 12
[Switch-LoopBack12] ip address 4.4.4.4 32
[Switch-LoopBack12] quit
# Specify IP address 4.4.4.4 as the listening IP address of the local portal server for Layer 2 portal authentication.
[Switch] portal local-server ip 4.4.4.4
# Enable portal authentication on port GigabitEthernet 1/0/1, and specify the Auth-Fail VLAN of the port as VLAN 2.
[Switch] interface gigabitethernet 1/0/1
[Switch-GigabitEthernet1/0/1] port link-type hybrid
[Switch-GigabitEthernet1/0/1] mac-vlan enable
[Switch-GigabitEthernet1/0/1] portal local-server enable
[Switch-GigabitEthernet1/0/1] portal auth-fail vlan 2
[Switch-GigabitEthernet1/0/1] quit

```

2. Configure a RADIUS scheme:

```

# Create a RADIUS scheme named rs1 and enter its view.
<Switch> system-view
[Switch] radius scheme rs1
# Set the server type for the RADIUS scheme. When using the IMC server, set the server type to extended.
[Switch-radius-rs1] server-type extended
# Specify the primary authentication server and primary accounting server, and configure the keys for communication with the servers.
[Switch-radius-rs1] primary authentication 1.1.1.2
[Switch-radius-rs1] primary accounting 1.1.1.2
[Switch-radius-rs1] key accounting radius
[Switch-radius-rs1] key authentication radius
[Switch-radius-rs1] quit

```

3. Configure an authentication domain:

```

# Create and enter ISP domain triple.
[Switch] domain triple
# Configure AAA methods for the ISP domain.
[Switch-isp-triple] authentication portal radius-scheme rs1
[Switch-isp-triple] authorization portal radius-scheme rs1
[Switch-isp-triple] accounting portal radius-scheme rs1
[Switch-isp-triple] quit
# Configure domain triple as the default ISP domain for all users. Then, if a user enters a username without any ISP domain at logon, the authentication and accounting methods of the default domain are used for the user.
[Switch] domain default enable triple

```

4. Configure the DHCP relay agent:

```

# Enable DHCP.
[Switch] dhcp enable

```



```

# Create DHCP server group 1 and add DHCP server 1.1.1.3 into the group.
[Switch] dhcp relay server-group 1 ip 1.1.1.3
# Enable the DHCP relay agent on VLAN-interface 8.
[Switch] interface vlan-interface 8
[Switch-Vlan-interface8] dhcp select relay
# Correlate DHCP server group 1 with VLAN-interface 8.
[Switch-Vlan-interface8] dhcp relay server-select 1
[Switch-Vlan-interface8] quit
# Enable the DHCP relay agent on VLAN-interface 2.
[Switch] interface vlan-interface 2
[Switch-Vlan-interface2] dhcp select relay
# Correlate DHCP server group 1 with VLAN-interface 2.
[Switch-Vlan-interface2] dhcp relay server-select 1
[Switch-Vlan-interface2] quit
# Enable the DHCP relay agent on VLAN-interface 3.
[Switch] interface vlan-interface 3
[Switch-Vlan-interface3] dhcp select relay
# Correlate DHCP server group 1 with VLAN-interface 3.
[Switch-Vlan-interface3] dhcp relay server-select 1
[Switch-Vlan-interface3] quit

```

Verifying the configuration

Before user **userpt** accesses a web page, the user is in VLAN 8 (the initial VLAN), and is assigned with an IP address on subnet 192.168.1.0/24. When the user accesses a web page on the external network, the web request will be redirected to authentication page **https://4.4.4.4/portal/logon.htm**. After entering the correct username and password, the user can pass the authentication. Then, the device will move the user from VLAN 8 to VLAN 3, the authorized VLAN. You can use the **display connection ucibindex** command to view the online user information

```

<Switch> display connection ucibindex 30
Slot: 1
Index=30 , Username=userpt@triple
MAC=0015-e9a6-7cfe
IP=192.168.1.2
IPv6=N/A
Access=PORTAL ,AuthMethod=PAP
Port Type=Ethernet,Port Name=GigabitEthernet1/0/1
Initial VLAN=8, Authorization VLAN=3
ACL Group=Disable
User Profile=N/A
CAR=Disable
Priority=Disable
Start=2009-11-26 17:40:02 ,Current=2009-11-26 17:48:21 ,Online=00h08m19s
Total 1 connection matched.

```

Use the **display mac-vlan all** command to view the generated MAC-VLAN entries, which record the MAC addresses passing authentication and the corresponding VLANs.

```

[Switch] display mac-vlan all
The following MAC VLAN addresses exist:

```

```

S:Static  D:Dynamic
MAC ADDR          MASK                VLAN ID  PRIO    STATE
-----
0015-e9a6-7cfe   ffff-ffff-ffff  3        0       D
Total MAC VLAN address count:1

```

If a client fails authentication, it is added to VLAN 2. Use the previously mentioned commands to view the assigned IP address and the generated MAC-VLAN entry for the client.

Troubleshooting portal

Inconsistent keys on the access device and the portal server

Symptom

When a user is forced to access the portal server, the portal server displays a blank web page, rather than the portal authentication page or an error message.

Analysis

The keys configured on the access device and the portal server are inconsistent, causing CHAP message exchange failure. As a result, the portal server does not display the authentication page.

Solution

- Use the **display portal server** command to display the key for the portal server on the access device and view the key for the access device on the portal server.
- Use the **portal server** command to modify the key on the access device or modify the key for the access device on the portal server to make sure that the keys are consistent.

Incorrect server port number on the access device

Symptom

After a user passes the portal authentication, you cannot force the user to log off by executing the **portal delete-user** command on the access device, but the user can log off by using the **disconnect** attribute on the authentication client.

Analysis

When you execute the **portal delete-user** command on the access device to force the user to log off, the access device actively sends a REQ_LOGOUT message to the portal server. The default listening port of the portal server is 50100. However, if the listening port configured on the access device is not 50100, the destination port of the REQ_LOGOUT message is not the actual listening port on the server, and the portal server cannot receive the REQ_LOGOUT message. As a result, you cannot force the user to log off the portal server.

When the user uses the **disconnect** attribute on the client to log off, the portal server actively sends a REQ_LOGOUT message to the access device. The source port is 50100 and the destination port of the ACK_LOGOUT message from the access device is the source port of the REQ_LOGOUT message so that the portal server can receive the ACK_LOGOUT message correctly, no matter whether the listening port is configured on the access device. The user can log off the portal server.

Solution

Use the **display portal server** command to display the listening port of the portal server configured on the access device and use the **portal server** command in the system view to modify it to make sure that it is the actual listening port of the portal server.

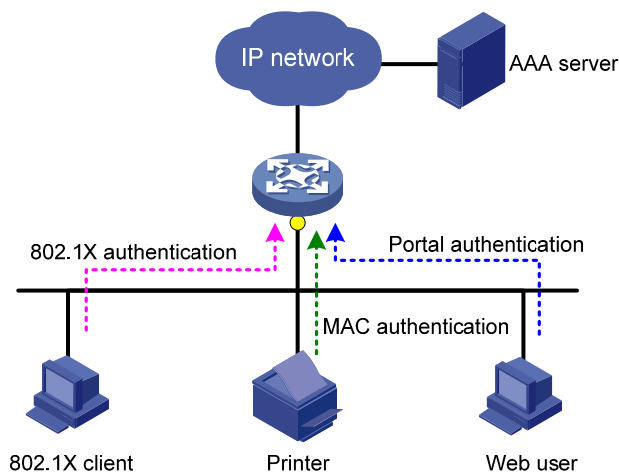
Configuring triple authentication

Overview

Triple authentication enables a Layer 2 access port to perform portal, MAC, and 802.1X authentication. A terminal can access the network if it passes one type of authentication.

Triple authentication is suitable for a LAN that comprises terminals that require different authentication services. For example, the triple authentication-enabled access port in [Figure 84](#) can perform MAC authentication for the printer, 802.1X authentication for a PC installed with the 802.1X client, and port authentication for the other PC.

Figure 84 Triple authentication network diagram



For more information about portal authentication, MAC authentication and 802.1X authentication, see "[Configuring portal authentication](#)," "[Configuring MAC authentication](#)," and "[Configuring 802.1X](#)."

Triple authentication mechanism

The three types of authentication are triggered by different packets:

- The access port performs MAC authentication for a terminal when it receives an ARP or DHCP broadcast packet from the terminal for the first time. If the terminal passes MAC authentication, the terminal can access the network. If the MAC authentication fails, the access port performs 802.1X or portal authentication.
- The access port performs 802.1X authentication when it receives an EAP packet from an 802.1X client. If the unicast trigger function of 802.1X is enabled on the access port, any packet from an 802.1X client can trigger an 802.1X authentication.
- The access port performs portal authentication when it receives an HTTP packet from a terminal.

If a terminal triggers different types of authentication, the authentications are processed at the same time. The failure of one type of authentication does not affect the others. When a terminal passes one type of authentication, the other types of authentication being performed are terminated. Then, whether the other types of authentication can be triggered varies:

- If a terminal passes 802.1X or portal authentication, no other types of authentication will be triggered for the terminal.
- If the terminal passes MAC authentication, no portal authentication can be triggered for the terminal, but 802.1X authentication can be triggered. When the terminal passes 802.1X authentication, the 802.1X authentication information will overwrite the MAC authentication information for the terminal.

Using triple authentication with other features

A triple authentication enabled access port supports working with the following features.

VLAN assignment

After a terminal passes authentication, the authentication server assigns an authorized VLAN to the access port for the access terminal. The terminal can then access the network resources in the authorized VLAN.

Auth-Fail VLAN or MAC authentication guest VLAN

After a terminal fails authentication, the access port:

- Adds the terminal to an Auth-Fail VLAN, if it uses 802.1X or portal authentication service.
- Adds the terminal to a MAC authentication guest VLAN, if it uses MAC authentication service.

A terminal may undergo all three types of authentication. If it fails to pass all types of authentication, the access port adds the terminal to the 802.1X Auth-Fail VLAN.

ACL assignment

You can specify an authorization ACL for an authenticated user to control its access to network resources. After the user passes MAC authentication, the authentication server, either the local access device or a RADIUS server, assigns the ACL onto the access port to filter traffic for the user.

You must configure the ACLs on the access device, whether the authentication server is the access device or a remote AAA server.

Detection of online terminals

- You can enable an online detection timer, which is configurable, to detect online portal clients.
- You can enable the online handshake or periodic re-authentication function to detect online 802.1X clients at a configurable interval.
- You can enable an offline detection timer to detect online MAC authentication terminals at a configurable interval.

For more information about the extended functions, see "[Configuring 802.1X](#)," "[Configuring MAC authentication](#)," and "[Configuring portal authentication](#)."

Configuring triple authentication

Step	Command	Remarks
1. Configure 802.1X authentication.	See " Configuring 802.1X "	Configure at least one type of authentication.
2. Configure MAC authentication.	See " Configuring MAC authentication "	802.1X authentication must use

Step	Command	Remarks
3. Configure Layer-2 portal authentication.	See " Configuring portal authentication "	MAC-based access control. HP does not recommend you configure 802.1X guest VLANs for triple authentication.

Triple authentication configuration examples

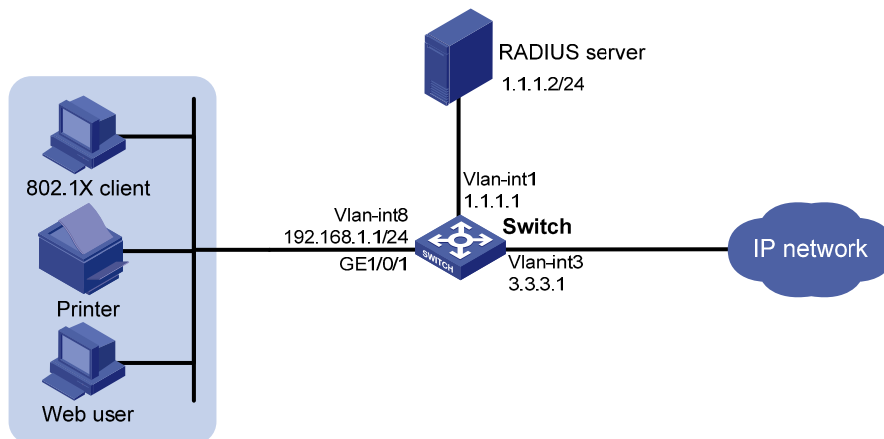
Triple authentication basic function configuration example

Network requirements

As shown in [Figure 85](#), the terminals are connected to a switch to access the IP network. Configure triple authentication on the Layer-2 interface of the switch that connects to the terminals so that a terminal passing one of the three authentication methods, 802.1X authentication, portal authentication, and MAC authentication, can access the IP network.

- Configure static IP addresses in network 192.168.1.0/24 for the terminals.
- Use the remote RADIUS server to perform authentication, authorization, and accounting and configure the switch to send usernames carrying no ISP domain names to the RADIUS server.
- The local portal authentication server on the switch uses listening IP address 4.4.4.4. The switch sends a default authentication page to the web user and forwards authentication data using HTTP.

Figure 85 Network diagram



Configuration procedure

Make sure that the terminals, the server, and the switch can reach each other.

The host of the web user must have a route to the listening IP address of the local portal server.

1. Configure the RADIUS server, and make sure the authentication, authorization, and accounting functions work normally. In this example, configure on the RADIUS server an 802.1X user (with username **userdot**), a portal user (with username **userpt**), and a MAC authentication user (with a username and password both being the MAC address of the printer **001588f80dd7**).
2. Configure portal authentication:
 - # Configure VLANs and IP addresses for the VLAN interfaces, and add ports to specific VLANs. (Details not shown.)

- ```
Configure the local portal server to support HTTP.
<Switch> system-view
[Switch] portal local-server http

Configure the IP address of interface loopback 0 as 4.4.4.4.
[Switch] interface loopback 0
[Switch-LoopBack0] ip address 4.4.4.4 32
[Switch-LoopBack0] quit

Specify the listening IP address of the local portal server for Layer-2 portal authentication as 4.4.4.4.
[Switch] portal local-server ip 4.4.4.4

Enable Layer-2 portal authentication on GigabitEthernet 1/0/1.
[Switch] interface gigabitethernet 1/0/1
[Switch-GigabitEthernet1/0/1] portal local-server enable
[Switch-GigabitEthernet1/0/1] quit
```
- 3. Configure 802.1X authentication:**
- ```
# Enable 802.1X authentication globally.
[Switch] dot1x

# Enable 802.1X authentication (MAC-based access control required) on GigabitEthernet 1/0/1.
[Switch] interface gigabitethernet 1/0/1
[Switch-GigabitEthernet1/0/1] dot1x port-method macbased
[Switch-GigabitEthernet1/0/1] dot1x
[Switch-GigabitEthernet1/0/1] quit
```
- 4. Configure MAC authentication:**
- ```
Enable MAC authentication globally.
[Switch] mac-authentication

Enable MAC authentication on GigabitEthernet 1/0/1.
[Switch] interface gigabitethernet 1/0/1
[Switch-GigabitEthernet1/0/1] mac-authentication
[Switch-GigabitEthernet1/0/1] quit
```
- 5. Configure a RADIUS scheme:**
- ```
# Create a RADIUS scheme named rs1.
[Switch] radius scheme rs1

# Specify the server type for the RADIUS scheme, which must be extended when the IMC server is used.
[Switch-radius-rs1] server-type extended

# Specify the primary authentication and accounting servers and keys.
[Switch-radius-rs1] primary authentication 1.1.1.2
[Switch-radius-rs1] primary accounting 1.1.1.2
[Switch-radius-rs1] key authentication radius
[Switch-radius-rs1] key accounting radius

# Specify usernames sent to the RADIUS server to carry no domain names.
[Switch-radius-rs1] user-name-format without-domain
[Switch-radius-rs1] quit
```
- 6. Configure an ISP domain:**
- ```
Create an ISP domain named triple.
```

```

[Switch] domain triple
Configure the default AAA methods for all types of users in the domain.
[Switch-isp-triple] authentication default radius-scheme rs1
[Switch-isp-triple] authorization default radius-scheme rs1
[Switch-isp-triple] accounting default radius-scheme rs1
[Switch-isp-triple] quit
Configure domain triple as the default domain. If a username input by a user includes no ISP
domain name, the authentication scheme of the default domain is used.
[Switch] domain default enable triple

```

## Verifying the configuration

User **userdot** uses the 802.1X client to initiate authentication. After inputting the correct username and password, the user can pass 802.1X authentication. Web user **userpt** uses a web browser to access an external network. The web request is redirected to the authentication page <http://4.4.4.4/portal/logon.htm>. After inputting the correct username and password, the web user can pass portal authentication. The printer can pass MAC authentication after being connected to the network.

Use the **display connection** command to view online users.

```

[Switch] display connection
Slot: 1
Index=30 , Username=userpt@triple
IP=192.168.1.2
IPv6=N/A
MAC=0015-e9a6-7cfe
Index=31 , Username=userdot@triple
IP=192.168.1.3
IPv6=N/A
MAC=0002-0002-0001
Index=32 , Username=001588f80dd7@triple
IP=192.168.1.4
IPv6=N/A
MAC=0015-88f8-0dd7

Total 3 connection(s) matched on slot 1.
Total 3 connection(s) matched.

```

## Triple authentication supporting VLAN assignment and Auth-Fail VLAN configuration example

### Network requirement

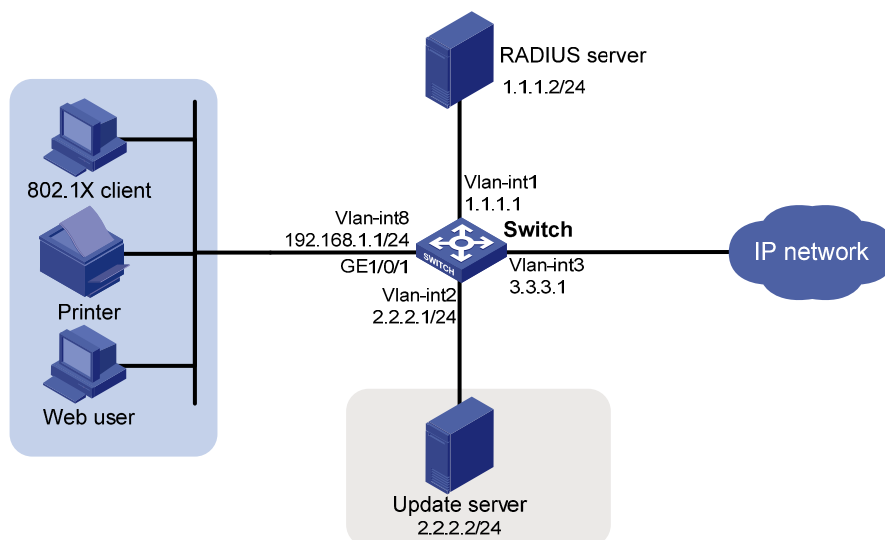
As shown in [Figure 86](#), the terminals are connected to a switch to access the IP network. Configure triple authentication on the Layer-2 interface of the switch which connects to the terminals so that a terminal passing one of the three authentication methods, 802.1X authentication, portal authentication, and MAC authentication, can access the IP network.

- Portal terminals use DHCP to get IP addresses in 192.168.1.0/24 before authentication and in 3.3.3.0/24 after passing authentication.



- 802.1X terminals use IP addresses in 192.168.1.0/24 before authentication, and request IP addresses in 3.3.3.0/24 through DHCP after passing authentication. If the terminal fails authentication, it uses an IP address in 2.2.2.0/24.
- After passing authentication, the printer obtains the IP address 3.3.3.111/24 that is bound with its MAC address through DHCP.
- Use the remote RADIUS server to perform authentication, authorization, and accounting and configure the switch to remove the ISP domain names from usernames sent to the RADIUS server.
- The local portal authentication server on the switch uses listening IP address 4.4.4.4. The switch sends a default authentication page to the web user and forwards authentication data by using HTTPS.
- Configure VLAN 3 as the authorized VLAN on the RADIUS server. Users passing authentication are added to this VLAN.
- Configure VLAN 2 as the Auth-Fail VLAN on the access device. Users failing authentication are added to this VLAN, and are allowed to access only the Update server.

**Figure 86 Network diagram**



## Configuration procedure

Make sure that the terminals, the servers, and the switch can reach each other.

When using an external DHCP server, make sure that the terminals can get IP addresses from the server before and after authentication.

1. Configure the RADIUS server, and make sure the authentication, authorization, and accounting functions work normally. In this example, configure on the RADIUS server an 802.1X user (with username **userdot**), a portal user (with username **userpt**), a MAC authentication user (with a username and password both being the MAC address of the printer **001588f80dd7**), and an authorized VLAN (VLAN 3).
2. Configure PKI domain **pkidm** and acquire the local and CA certificates. For more information, see "[Configuring PKI.](#)"
3. Complete the editing of a self-defined default authentication page file, compress the file to a zip file named defaultfile and save the zip file at the root directory.
4. Configure DHCP:

# Configure VLANs and IP addresses for the VLAN interfaces, and add ports to specific VLANs. (Details not shown.)

# Enable DHCP.

```
<Switch> system-view
[Switch] dhcp enable
```

# Exclude the IP address of the update server from assignment.

```
[Switch] dhcp server forbidden-ip 2.2.2.2
```

# Configure IP address pool 1, including the address range, lease and gateway address. A short lease is recommended to shorten the time terminals use to re-acquire IP addresses after the terminals passing or failing authentication.

```
[Switch] dhcp server ip-pool 1
[Switch-dhcp-pool-1] network 192.168.1.0 mask 255.255.255.0
[Switch-dhcp-pool-1] expired day 0 hour 0 minute 1
[Switch-dhcp-pool-1] gateway-list 192.168.1.1
[Switch-dhcp-pool-1] quit
```

A short lease is recommended to shorten the time that terminals use to re-acquire IP addresses after passing or failing authentication. However, in some applications, a terminal can require a new IP address before the lease duration expires. For example, the iNode 802.1X client automatically renews its IP address after disconnecting from the server.

# Configure IP address pool 2, including the address range, lease and gateway address. A short lease is recommended to shorten the time terminals use to re-acquire IP addresses after the terminals pass authentication.

```
[Switch] dhcp server ip-pool 2
[Switch-dhcp-pool-2] network 2.2.2.0 mask 255.255.255.0
[Switch-dhcp-pool-2] expired day 0 hour 0 minute 1
[Switch-dhcp-pool-2] gateway-list 2.2.2.1
[Switch-dhcp-pool-2] quit
```

# Configure IP address pool 3, including the address range, lease and gateway address. A short lease is recommended to shorten the time terminals use to re-acquire IP addresses after the terminals are offline.

```
[Switch] dhcp server ip-pool 3
[Switch-dhcp-pool-3] network 3.3.3.0 mask 255.255.255.0
[Switch-dhcp-pool-3] expired day 0 hour 0 minute 1
[Switch-dhcp-pool-3] gateway-list 3.3.3.1
[Switch-dhcp-pool-3] quit
```

# Configure IP address pool 4, and bind the printer MAC address 0015-e9a6-7cfe to the IP address 3.3.3.111/24 in this address pool.

```
[Switch] dhcp server ip-pool 4
[Switch-dhcp-pool-4] static-bind ip-address 3.3.3.111 mask 255.255.255.0
[Switch-dhcp-pool-4] static-bind mac-address 0015-e9a6-7cfe
[Switch-dhcp-pool-4] quit
```

## 5. Configure portal authentication:

# Create SSL server policy **sslsvr** and specify it to use PKI domain **pkidm**.

```
[Switch] ssl server-policy sslsvr
[Switch-ssl-server-policy-sslsvr] pki pkidm
[Switch-ssl-server-policy-sslsvr] quit
```

# Configure the local portal server to support HTTPS and use SSL server policy **sslsvr**.

```
[Switch] portal local-server https server-policy sslsvr
Configure IP address 4.4.4.4 for interface loopback 12.
[Switch] interface loopback 12
[Switch-LoopBack12] ip address 4.4.4.4 32
[Switch-LoopBack12] quit
Specify the listening IP address of the local portal server as 4.4.4.4.
[Switch] portal local-server ip 4.4.4.4
Enable Layer-2 portal authentication on GigabitEthernet 1/0/1 and specify VLAN 2 as the
Auth-Fail VLAN, to which terminals failing authentication are added.
[Switch] interface gigabitethernet 1/0/1
[Switch-GigabitEthernet1/0/1] port link-type hybrid
[Switch-GigabitEthernet1/0/1] mac-vlan enable
[Switch-GigabitEthernet1/0/1] portal local-server enable
[Switch-GigabitEthernet1/0/1] portal auth-fail vlan 2
[Switch-GigabitEthernet1/0/1] quit
```

#### 6. Configure 802.1X authentication:

# Enable 802.1X authentication globally.

```
[Switch] dot1x
```

# Enable 802.1X authentication (MAC-based access control required) on GigabitEthernet 1/0/1, and specify VLAN 2 as the Auth-Fail VLAN.

```
[Switch] interface gigabitethernet 1/0/1
[Switch-GigabitEthernet1/0/1] dot1x port-method macbased
[Switch-GigabitEthernet1/0/1] dot1x
[Switch-GigabitEthernet1/0/1] dot1x auth-fail vlan 2
[Switch-GigabitEthernet1/0/1] quit
```

#### 7. Configure MAC authentication:

# Enable MAC authentication globally.

```
[Switch] mac-authentication
```

# Enable MAC authentication on GigabitEthernet 1/0/1, and specify VLAN 2 as the Auth-Fail VLAN

```
[Switch] interface gigabitethernet 1/0/1
[Switch-GigabitEthernet1/0/1] mac-authentication
[Switch-GigabitEthernet1/0/1] mac-authentication guest-vlan 2
[Switch-GigabitEthernet1/0/1] quit
```

#### 8. Configure a RADIUS scheme:

# Create a RADIUS scheme named **rs1**.

```
[Switch] radius scheme rs1
```

# Specify the server type for the RADIUS scheme, which must be **extended** when the IMC server is used.

```
[Switch-radius-rs1] server-type extended
```

# Specify the primary authentication and accounting servers and keys.

```
[Switch-radius-rs1] primary authentication 1.1.1.2
```

```
[Switch-radius-rs1] primary accounting 1.1.1.2
```

```
[Switch-radius-rs1] key authentication radius
```

```
[Switch-radius-rs1] key accounting radius
```

# Specify usernames sent to the RADIUS server to carry no domain names.

```
[Switch-radius-rs1] user-name-format without-domain
[Switch-radius-rs1] quit
```

## 9. Configure an ISP domain:

# Create an ISP domain named triple.

```
[Switch] domain triple
```

# Configure the default AAA methods for all types of users in the domain.

```
[Switch-isp-triple] authentication default radius-scheme rs1
```

```
[Switch-isp-triple] authorization default radius-scheme rs1
```

```
[Switch-isp-triple] accounting default radius-scheme rs1
```

```
[Switch-isp-triple] quit
```

# Configure domain **triple** as the default domain. If a username input by a user includes no ISP domain name, the authentication scheme of the default domain is used.

```
[Switch] domain default enable triple
```

## Verifying the configuration

User **userdot** uses the 802.1X client to initiate authentication. After inputting the correct username and password, the user can pass 802.1X authentication. Web user **userpt** uses a web browser to access an external network. The web request is redirected to the authentication page <http://4.4.4.4/portal/logon.htm>. After inputting the correct username and password, the web user can pass portal authentication. The printer can pass MAC authentication after being connected to the network.

Use the **display connection** command to view connection information about online users.

```
[Switch] display connection
Slot: 1
Index=30 , Username=userpt@triple
IP=192.168.1.2
IPv6=N/A
MAC=0015-e9a6-7cfe
Index=31 , Username=userdot@triple
IP=3.3.3.2
IPv6=N/A
MAC=0002-0002-0001
Index=32 , Username=001588f80dd7@triple
IP=N/A
IPv6=N/A
MAC=0015-88f8-0dd7
```

Total 3 connection(s) matched on slot 1.

Total 3 connection(s) matched.

Use the **display mac-vlan all** command to view the MAC-VLAN entries of online users. VLAN 3 is the authorized VLAN.

```
[Switch] display mac-vlan all
The following MAC VLAN addresses exist:
S:Static D:Dynamic
MAC ADDR MASK VLAN ID PRIO STATE

0015-e9a6-7cfe ffff-ffff-ffff 3 0 D
```

```
0002-0002-0001 ffff-ffff-ffff 3 0 D
0015-88f8-0dd7 ffff-ffff-ffff 3 0 D
Total MAC VLAN address count:3
```

Use the **display dhcp server ip-in-use** command to view the IP addresses assigned to online users.

```
[Switch] display dhcp server ip-in-use all
```

```
Pool utilization: 0.59%
```

| IP address | Client-identifier/<br>Hardware address | Lease expiration     | Type           |
|------------|----------------------------------------|----------------------|----------------|
| 3.3.3.111  | 0015-88f8-0dd7                         | Dec 15 2009 17:40:52 | Auto:COMMITTED |
| 3.3.3.2    | 0002-0002-0001                         | Dec 15 2009 17:41:02 | Auto:COMMITTED |
| 3.3.3.3    | 0015-e9a6-7cfe                         | Unlimited            | Manual         |

```
--- total 3 entry ---
```

When a terminal fails authentication, it is added to VLAN 2. You can also use the display commands to view the MAC-VLAN entry and IP address of the terminal.

---

# Configuring port security

## Overview

Port security combines and extends 802.1X and MAC authentication to provide MAC-based network access control. It applies to a network that requires different authentication methods for different users on a port.

Port security prevents unauthorized access to the network by checking the source MAC address of inbound traffic and prevents access to unauthorized devices by checking the destination MAC address of outbound traffic.

Port security can control MAC address learning and authentication on a port to make sure that the port learns only trusted MAC addresses.

A frame is illegal, if its source MAC address cannot be learned in a port security mode or it is from a client that has failed 802.1X or MAC authentication.

The port security feature can automatically take a pre-defined action on illegal frames. This automatic mechanism enhances network security and reduces human intervention.

---

### NOTE:

For scenarios that require only 802.1X authentication or MAC authentication, HP recommends you configure 802.1X authentication or MAC authentication rather than port security. For more information about 802.1X and MAC authentication, see "[Configuring 802.1X](#)" and "[Configuring MAC authentication](#)."

---

## Port security features

### NTK

The need to know (NTK) feature prevents traffic interception by checking the destination MAC address in the outbound frames. The feature guarantees that frames are sent only to hosts that have passed authentication or whose MAC addresses have been learned or configured on the access device.

### Intrusion protection

The intrusion protection feature checks the source MAC address in inbound frames for illegal frames and takes a pre-defined action on each detected illegal frame. The action can be disabling the port temporarily, disabling the port permanently, or blocking frames from the illegal MAC address for three minutes (not user configurable).

### Port security traps

You can configure the port security module to send traps for port security events such as login, logoff, and MAC authentication. These traps help you monitor user behaviors.

## Port security modes

Port security supports the following categories of security modes:

- **MAC learning control**—Includes two modes, autoLearn and secure. MAC address learning is permitted on a port in autoLearn mode and disabled in secure mode.
- **Authentication**—Security modes in this category implement MAC authentication, 802.1X authentication, or a combination of these two authentication methods.

Upon receiving a frame, the port in a security mode searches the MAC address table for the source MAC address. If a match is found, the port forwards the frame. If no match is found, the port learns the MAC address or performs authentication, depending on the security mode. If the frame is illegal, the port takes the pre-defined NTK, intrusion protection, or trapping action.

The maximum number of users a port supports equals the maximum number of MAC addresses that port security allows or the maximum number of concurrent users the authentication mode in use allows, whichever is smaller. For example, if 802.1X allows more concurrent users than port security's limit on the number of MAC addresses on the port in userLoginSecureExt mode, port security's limit takes effect.

Table 11 describes the port security modes and the security features.

**Table 11 Port security modes**

| Purpose                                                                  | Security mode                                                                                                                      | Features that can be triggered   |                          |
|--------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------|----------------------------------|--------------------------|
| Turning off the port security feature                                    | noRestrictions (the default mode)<br>In this mode, port security is disabled on the port and access to the port is not restricted. | N/A                              |                          |
| Controlling MAC address learning                                         | autoLearn                                                                                                                          | NTK/intrusion protection         |                          |
|                                                                          | secure                                                                                                                             |                                  |                          |
| Performing 802.1X authentication                                         | userLogin                                                                                                                          | N/A                              |                          |
|                                                                          | userLoginSecure                                                                                                                    |                                  |                          |
|                                                                          | userLoginSecureExt                                                                                                                 |                                  |                          |
|                                                                          | userLoginWithOUI                                                                                                                   |                                  |                          |
| Performing MAC authentication                                            | macAddressWithRadius                                                                                                               | NTK/intrusion protection         |                          |
| Performing a combination of MAC authentication and 802.1X authentication | Or                                                                                                                                 | macAddressOrUserLoginSecure      | NTK/intrusion protection |
|                                                                          |                                                                                                                                    | macAddressOrUserLoginSecureExt   |                          |
|                                                                          | Else                                                                                                                               | macAddressElseUserLoginSecure    |                          |
|                                                                          |                                                                                                                                    | macAddressElseUserLoginSecureExt |                          |



**TIP:**

- **userLogin** specifies 802.1X authentication and port-based access control.
- **macAddress** specifies MAC authentication.
- **Else** specifies that the authentication method before **Else** is applied first. If the authentication fails, whether to turn to the authentication method following **Else** depends on the protocol type of the authentication request.
- Typically, in a security mode with **Or**, the authentication method to be used depends on the protocol type of the authentication request.
- **userLogin** with **Secure** specifies 802.1X authentication and MAC-based access control.
- **Ext** indicates allowing multiple 802.1X users to be authenticated and serviced at the same time. A security mode without **Ext** allows only one user to pass 802.1X authentication.

## Controlling MAC address learning

- autoLearn

A port in this mode can learn MAC addresses, and allows frames from learned or configured MAC addresses to pass. The automatically learned MAC addresses are secure MAC addresses. You can also configure secure MAC addresses by using the **port-security mac-address security** command. A secure MAC address never ages out by default.

When the number of secure MAC addresses reaches the upper limit, the port transitions to secure mode.

The dynamic MAC address learning function in MAC address management is disabled on ports operating in autoLearn mode, but you can configure MAC addresses by using the **mac-address dynamic** and **mac-address static** commands.

- secure

MAC address learning is disabled on a port in secure mode. You configure MAC addresses by using the **mac-address static** and **mac-address dynamic** commands. For more information about configuring MAC address table entries, see *Layer 2—LAN Switching Configuration Guide*.

A port in secure mode allows only frames sourced from secure MAC addresses and manually configured MAC addresses to pass.

## Performing 802.1X authentication

- userLogin

A port in this mode performs 802.1X authentication and implements port-based access control. The port can service multiple 802.1X users. If one 802.1X user passes authentication, all the other 802.1X users of the port can access the network without authentication.

- userLoginSecure

A port in this mode performs 802.1X authentication and implements MAC-based access control. The port services only one user passing 802.1X authentication.

- userLoginSecureExt

This mode is similar to the userLoginSecure mode except that this mode supports multiple online 802.1X users.

- userLoginWithOUI

This mode is similar to the userLoginSecure mode. The difference is that a port in this mode also permits frames from one user whose MAC address contains a specific organizationally unique identifier (OUI).

For wired users, the port performs 802.1X authentication upon receiving 802.1X frames, and performs OUI check upon receiving non-802.1X frames.

## Performing MAC authentication

macAddressWithRadius: A port in this mode performs MAC authentication and services multiple users.

## Performing a combination of MAC authentication and 802.1X authentication

- macAddressOrUserLoginSecure

This mode is the combination of the macAddressWithRadius and userLoginSecure modes.

For wired users, the port performs MAC authentication upon receiving non-802.1X frames and performs 802.1X authentication upon receiving 802.1X frames.

- macAddressOrUserLoginSecureExt



This mode is similar to the `macAddressOrUserLoginSecure` mode except that a port in this mode supports multiple 802.1X and MAC authentication users.

- `macAddressElseUserLoginSecure`

This mode is the combination of the `macAddressWithRadius` and `userLoginSecure` modes, with MAC authentication having a higher priority as the **Else** keyword implies.

For non-802.1X frames, a port in this mode performs only MAC authentication. For 802.1X frames, it performs MAC authentication and then, if the authentication fails, 802.1X authentication.

- `macAddressElseUserLoginSecureExt`

This mode is similar to the `macAddressElseUserLoginSecure` mode except that a port in this mode supports multiple 802.1X and MAC authentication users as the keyword **Ext** implies.

#### NOTE:

An OUI, as defined by the IEEE, is the first 24 bits of the MAC address, which uniquely identifies a device vendor.

## Working with guest VLAN and Auth-Fail VLAN

An 802.1X guest VLAN is the VLAN that a user is in before initiating authentication. An 802.1X Auth-Fail VLAN or a MAC authentication guest VLAN is the VLAN that a user is in after failing authentication. Support for the guest VLAN and Auth-Fail VLAN features varies with security modes.

- You can use the 802.1X guest VLAN and 802.1X Auth-Fail VLAN features together with port security modes that support 802.1X authentication. For more information about the 802.1X guest VLAN and Auth-Fail VLAN on a port that performs MAC-based access control, see "[Configuring 802.1X](#)."
- You can use the MAC authentication VLAN feature together with security modes that support MAC authentication. For more information about the MAC authentication guest VLAN, see "[Configuring MAC authentication](#)."
- If you configure both an 802.1X Auth-Fail VLAN and a MAC authentication guest VLAN on a port that performs MAC-based access control, the 802.1X Auth-Fail VLAN has a higher priority.

## Configuration task list

| Task                                                                                   | Remarks                                          |                                             |
|----------------------------------------------------------------------------------------|--------------------------------------------------|---------------------------------------------|
| <a href="#">Enabling port security</a>                                                 | Required.                                        |                                             |
| <a href="#">Setting port security's limit on the number of MAC addresses on a port</a> | Optional.                                        |                                             |
| <a href="#">Setting the port security mode</a>                                         | Required.                                        |                                             |
| <a href="#">Configuring port security features</a>                                     | <a href="#">Configuring NTK</a>                  | Optional.                                   |
|                                                                                        | <a href="#">Configuring intrusion protection</a> | Configure one or more features as required. |
|                                                                                        | <a href="#">Enabling port security traps</a>     |                                             |
| <a href="#">Configuring secure MAC addresses</a>                                       | Optional.                                        |                                             |
| <a href="#">Ignoring authorization information from the server</a>                     | Optional.                                        |                                             |

## Enabling port security

Enabling or disabling port security resets the following security settings to the default:

- 802.1X access control mode is MAC-based, and the port authorization state is auto.
- Port security mode is noRestrictions.

When port security is enabled, you cannot manually enable 802.1X or MAC authentication, or change the access control mode or port authorization state. The port security automatically modifies these settings in different security modes.

You cannot disable port security when online users are present.

Before enabling port security, disable 802.1X and MAC authentication globally.

To enable port security:

| Step                     | Command                     | Remarks                                    |
|--------------------------|-----------------------------|--------------------------------------------|
| 1. Enter system view.    | <b>system-view</b>          | N/A                                        |
| 2. Enable port security. | <b>port-security enable</b> | By default, the port security is disabled. |

For more information about 802.1X configuration, see "[Configuring 802.1X](#)." For more information about MAC authentication configuration, see "[Configuring MAC authentication](#)."

## Setting port security's limit on the number of MAC addresses on a port

You can set the maximum number of MAC addresses that port security allows on a port for the following purposes:

- Controlling the number of concurrent users on the port. The maximum number of concurrent users on the port equals this limit or the limit of the authentication mode (802.1X for example) in use, whichever is smaller.
- Controlling the number of secure MAC addresses on the port in autoLearn mode.

The port security's limit on the number of MAC addresses on a port is independent of the MAC learning limit described in MAC address table configuration in the *Layer 2—LAN Switching Configuration Guide*.

To set the maximum number of secure MAC addresses allowed on a port:

| Step                                                              | Command                                                           | Remarks                 |
|-------------------------------------------------------------------|-------------------------------------------------------------------|-------------------------|
| 1. Enter system view.                                             | <b>system-view</b>                                                | N/A                     |
| 2. Enter Layer 2 Ethernet interface view.                         | <b>interface</b> <i>interface-type</i><br><i>interface-number</i> | N/A                     |
| 3. Set the limit of port security on the number of MAC addresses. | <b>port-security max-mac-count</b><br><i>count-value</i>          | Not limited by default. |

# Setting the port security mode

After enabling port security, you can change the port security mode of a port only when the port is operating in noRestrictions (the default) mode. To change the port security mode for a port in any other mode, first use the **undo port-security port-mode** command to restore the default port security mode.

You can specify a port security mode when port security is disabled, but your configuration cannot take effect.

You cannot change the port security mode of a port when online users are present.

## Configuration prerequisites

Before you set a port security mode for a port, complete the following tasks:

- Disable 802.1X and MAC authentication.
- Verify that the port does not belong to any aggregation group or service loopback group.
- If you are configuring the autoLearn mode, set port security's limit on the number of MAC addresses. You cannot change the setting when the port is operating in autoLearn mode.

## Configuration procedure

To enable a port security mode:

| Step                                         | Command                                                                                                                                                                                                                                                                                                                                               | Remarks                                                                                                                        |
|----------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                        | <b>system-view</b>                                                                                                                                                                                                                                                                                                                                    | N/A                                                                                                                            |
| 2. Set an OUI value for user authentication. | <b>port-security oui <i>oui-value</i> index <i>index-value</i></b>                                                                                                                                                                                                                                                                                    | Required for the <b>userlogin-withoui</b> mode.<br>Not configured by default.<br>To set multiple OUI values, repeat this step. |
| 3. Enter Layer 2 Ethernet interface view.    | <b>interface <i>interface-type</i> <i>interface-number</i></b>                                                                                                                                                                                                                                                                                        | N/A                                                                                                                            |
| 4. Set the port security mode.               | <b>port-security port-mode { <b>autolearn</b>   <b>mac-authentication</b>   <b>mac-else-userlogin-secure</b>   <b>mac-else-userlogin-secure-ext</b>   <b>secure</b>   <b>userlogin</b>   <b>userlogin-secure</b>   <b>userlogin-secure-ext</b>   <b>userlogin-secure-or-mac</b>   <b>userlogin-secure-or-mac-ext</b>   <b>userlogin-withoui</b> }</b> | By default, a port operates in noRestrictions mode.                                                                            |

# Configuring port security features

## Configuring NTK

The NTK feature checks the destination MAC addresses in outbound frames to make sure that frames are forwarded only to authenticated devices. Any unicast frame with an unknown destination MAC address is discarded. Not all port security modes support triggering the NTK feature. For more information, see [Table 11](#).

The NTK feature supports the following modes:

- **ntkonly**—Forwards only unicast frames with authenticated destination MAC addresses.
- **ntk-withbroadcasts**—Forwards only broadcast frames and unicast frames with authenticated destination MAC addresses.
- **ntk-withmulticasts**—Forwards only broadcast frames, multicast frames, and unicast frames with authenticated destination MAC addresses.

To configure the NTK feature:

| Step                                      | Command                                                                                                        | Remarks                                                                      |
|-------------------------------------------|----------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------|
| 1. Enter system view.                     | <b>system-view</b>                                                                                             | N/A                                                                          |
| 2. Enter Layer 2 Ethernet interface view. | <b>interface</b> <i>interface-type</i><br><i>interface-number</i>                                              | N/A                                                                          |
| 3. Configure the NTK feature.             | <b>port-security ntk-mode</b><br>{ <b>ntk-withbroadcasts</b>  <br><b>ntk-withmulticasts</b>   <b>ntkonly</b> } | By default, NTK is disabled on a port and all frames are allowed to be sent. |

## Configuring intrusion protection

Intrusion protection enables a device to take one of the following actions in response to illegal frames:

- **blockmac**—Adds the source MAC addresses of illegal frames to the blocked MAC addresses list and discards the frames. All subsequent frames sourced from a blocked MAC address will be dropped. A blocked MAC address is restored to normal state after being blocked for three minutes. The interval is fixed and cannot be changed.
- **disableport**—Disables the port until you bring it up manually.
- **disableport-temporarily**—Disables the port for a specific period of time. The period can be configured with the **port-security timer disableport** command.

On a port operating in either the `macAddressElseUserLoginSecure` mode or the `macAddressElseUserLoginSecureExt` mode, intrusion protection is triggered only after both MAC authentication and 802.1X authentication for the same frame fail.

To configure the intrusion protection feature:

| Step                  | Command            | Remarks |
|-----------------------|--------------------|---------|
| 1. Enter system view. | <b>system-view</b> | N/A     |

| Step                                                                    | Command                                                                                                             | Remarks                                       |
|-------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------|-----------------------------------------------|
| 2. Enter Layer 2 Ethernet interface view.                               | <b>interface</b> <i>interface-type</i><br><i>interface-number</i>                                                   | N/A                                           |
| 3. Configure the intrusion protection feature.                          | <b>port-security intrusion-mode</b><br>{ <b>blockmac</b>   <b>disableport</b>  <br><b>disableport-temporarily</b> } | By default, intrusion protection is disabled. |
| 4. Return to system view.                                               | <b>quit</b>                                                                                                         | N/A                                           |
| 5. Set the silence timeout period during which a port remains disabled. | <b>port-security timer disableport</b><br><i>time-value</i>                                                         | Optional.<br>20 seconds by default.           |

## Enabling port security traps

You can configure the port security module to send traps for the following categories of events:

- **addresslearned**—Learning of new MAC addresses.
- **dot1xlogfailure/dot1xlogon/dot1xlogoff**—802.1X authentication failure, success, and 802.1X user logoff.
- **ralmlogfailure/ralmlogon/ralmlogoff**—MAC authentication failure, MAC authentication user logon, and MAC authentication user logoff.
- **intrusion**—Detection of illegal frames.

To enable port security traps:

| Step                           | Command                                                                                                                                                                                                             | Remarks                                       |
|--------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------|
| 1. Enter system view.          | <b>system-view</b>                                                                                                                                                                                                  | N/A                                           |
| 2. Enable port security traps. | <b>port-security trap</b> { <b>addresslearned</b><br>  <b>dot1xlogfailure</b>   <b>dot1xlogoff</b>  <br><b>dot1xlogon</b>   <b>intrusion</b>  <br><b>ralmlogfailure</b>   <b>ralmlogoff</b>  <br><b>ralmlogon</b> } | By default, port security traps are disabled. |

## Configuring secure MAC addresses

Secure MAC addresses are configured or learned in autoLearn mode and can survive link down/up events. You can bind a secure MAC address to only one port in a VLAN.

### ⓘ IMPORTANT:

When the maximum number of secure MAC address entries is reached, the port changes to secure mode, and no more secure MAC addresses can be added or learned. The port allows only frames sourced from a secure MAC address or a MAC address configured by using the **mac-address dynamic** or **mac-address static** command to pass through.

Secure MAC addresses fall into static, sticky and dynamic secure MAC addresses.

**Table 12 A comparison of static, sticky, and dynamic secure MAC addresses**

| Type    | Address sources                                                                                                                        | Aging mechanism                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | Can be saved and survive a device reboot?                   |
|---------|----------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------|
| Static  | Manually added                                                                                                                         | Not available.<br>They never age out unless you manually remove them, change the port security mode, or disable the port security feature.                                                                                                                                                                                                                                                                                                                                                                                                                        | Yes.                                                        |
| Sticky  | Manually added or automatically learned when the dynamic secure MAC function ( <b>port-security mac-address dynamic</b> ) is disabled. | Sticky MAC addresses by default do not age out, but you can configure an aging timer or use the aging timer together with the inactivity aging function to delete old sticky MAC addresses: <ul style="list-style-type: none"> <li>If only an aging timer is configured, the aging timer counts up regardless of whether traffic data has been sent from the sticky MAC address.</li> <li>If both an aging timer and the inactivity aging function are configured, the aging timer restarts once traffic data is detected from the sticky MAC address.</li> </ul> | Yes.<br>The secure MAC aging timer restarts at a reboot.    |
| Dynamic | Converted from sticky MAC addresses or automatically learned after the dynamic secure MAC function is enabled.                         | Same as sticky MAC addresses.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | No.<br>All dynamic secure MAC addresses are lost at reboot. |

## Configuration prerequisites

- Enable port security.
- Set port security's limit on the number of MAC addresses on the port. Perform this task before you enable autoLearn mode.
- Set the port security mode to autoLearn.

## Configuration procedure

To configure a secure MAC address:

| Step                               | Command                                               | Remarks                                                                                                                                                                                                                                       |
|------------------------------------|-------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.              | <b>system-view</b>                                    | N/A                                                                                                                                                                                                                                           |
| 2. Set the secure MAC aging timer. | <b>port-security timer autolearn aging time-value</b> | Optional.<br>By default, secure MAC addresses do not age out, and you can remove them only by performing the <b>undo port-security mac-address security</b> command, changing the port security mode, or disabling the port security feature. |

| Step                                       | Command                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | Remarks                                                                                                                            |
|--------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------|
| 3. Configure a secure MAC address.         | <ul style="list-style-type: none"> <li>• <b>Approach 1 (in system view):</b><br/> <b>port-security mac-address security</b> [ sticky] <i>mac-address</i><br/> <b>interface</b> <i>interface-type</i><br/> <i>interface-number</i> <b>vlan</b> <i>vlan-id</i></li> <li>• <b>Approach 2 (in interface view):</b> <ul style="list-style-type: none"> <li>a. <b>interface</b> <i>interface-type</i><br/> <i>interface-number</i></li> <li>b. <b>port-security mac-address security</b> [ sticky]<br/> <i>mac-address</i> <b>vlan</b> <i>vlan-id</i></li> <li>c. <b>quit</b></li> </ul> </li> </ul> | Use either approach.<br>No secure MAC address exists by default.                                                                   |
| 4. Enter Layer 2 Ethernet interface view.  | <b>interface</b> <i>interface-type</i><br><i>interface-number</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | N/A                                                                                                                                |
| 5. Enable inactivity aging.                | <b>port-security mac-address aging-type inactivity</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | Optional.<br>By default, the inactivity aging function is disabled.                                                                |
| 6. Enable the dynamic secure MAC function. | <b>port-security mac-address dynamic</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | Optional.<br>By default, sticky MAC addresses can be saved to the configuration file, and once saved, can survive a device reboot. |

**NOTE:**

You can display dynamic secure MAC addresses only by using the **display port-security mac-address security** command.

## Ignoring authorization information from the server

The authorization information is delivered by the RADIUS server to the device after an 802.1X user or MAC authenticated user passes RADIUS authentication. You can configure a port to ignore the authorization information from the RADIUS server.

To configure a port to ignore the authorization information from the RADIUS server:

| Step                                                            | Command                                                           | Remarks                                                                       |
|-----------------------------------------------------------------|-------------------------------------------------------------------|-------------------------------------------------------------------------------|
| 1. Enter system view.                                           | <b>system-view</b>                                                | N/A                                                                           |
| 2. Enter Layer 2 Ethernet interface view.                       | <b>interface</b> <i>interface-type</i><br><i>interface-number</i> | N/A                                                                           |
| 3. Ignore the authorization information from the RADIUS server. | <b>port-security authorization ignore</b>                         | By default, a port uses the authorization information from the RADIUS server. |

## Displaying and maintaining port security

| Task                                                                                                                         | Command                                                                                                                                                                                                                                                 | Remarks               |
|------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Display port security configuration information, operation information, and statistics about one or more ports or all ports. | <b>display port-security</b> [ <b>interface</b> <i>interface-list</i> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                                                                                              | Available in any view |
| Display information about secure MAC addresses.                                                                              | <b>display port-security mac-address security</b> [ <b>interface</b> <i>interface-type</i> <i>interface-number</i> ] [ <b>vlan</b> <i>vlan-id</i> ] [ <b>count</b> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] | Available in any view |
| Display information about blocked MAC addresses.                                                                             | <b>display port-security mac-address block</b> [ <b>interface</b> <i>interface-type</i> <i>interface-number</i> ] [ <b>vlan</b> <i>vlan-id</i> ] [ <b>count</b> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]    | Available in any view |

## Port security configuration examples

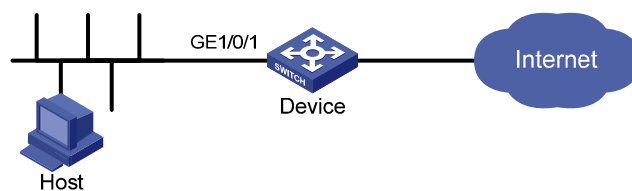
### Configuring the autoLearn mode

#### Network requirements

See [Figure 87](#). Configure port GigabitEthernet 1/0/1 on the Device, as follows:

- Accept up to 64 users on the port without authentication.
- Permit the port to learn and add MAC addresses as sticky MAC addresses, and set the sticky MAC aging timer to 30 minutes.
- After the number of secure MAC addresses reaches 64, the port stops learning MAC addresses. If any frame with an unknown MAC address arrives, intrusion protection starts, and the port shuts down and stays silent for 30 seconds.

**Figure 87 Network diagram**



#### Configuration procedure

```
Enable port security.
<Device> system-view
[Device] port-security enable

Set the secure MAC aging timer to 30 minutes.
[Device] port-security timer autolearn aging 30

Enable intrusion protection traps on port GigabitEthernet 1/0/1.
[Device] port-security trap intrusion
[Device] interface gigabitethernet 1/0/1
```



```

Set port security's limit on the number of MAC addresses to 64 on the port.
[Device-GigabitEthernet1/0/1] port-security max-mac-count 64

Set the port security mode to autoLearn.
[Device-GigabitEthernet1/0/1] port-security port-mode autolearn

Configure the port to be silent for 30 seconds after the intrusion protection feature is triggered.
[Device-GigabitEthernet1/0/1] port-security intrusion-mode disableport-temporarily
[Device-GigabitEthernet1/0/1] quit
[Device] port-security timer disableport 30

```

## Verifying the configuration

```

Display the port security configuration.
<Device> display port-security interface gigabitethernet 1/0/1
 Equipment port-security is enabled
 Intrusion trap is enabled
AutoLearn aging time is 30 minutes
 Disableport Timeout: 30s
 OUI value:

GigabitEthernet1/0/1 is link-up
 Port mode is autoLearn
 NeedToKnow mode is disabled
 Intrusion Protection mode is DisablePortTemporarily
 Max MAC address number is 64
 Stored MAC address number is 0
 Authorization is permitted
 Security MAC address learning mode is sticky
 Security MAC address aging type is absolute

```

The output shows that the port security's limit on the number of secure MAC addresses on the port is 64, the port security mode is autoLearn, intrusion protection traps are enabled, and the intrusion protection action is disabling the port (DisablePortTemporarily) for 30 seconds.

# Repeatedly perform the **display port-security** command to track the number of MAC addresses learned by the port, or use the **display this** command in Layer 2 Ethernet interface view to display the secure MAC addresses.

```

<Device> system-view
[Device] interface gigabitethernet 1/0/1
[Device-GigabitEthernet1/0/1] display this
#
interface GigabitEthernet1/0/1
 port-security max-mac-count 64
 port-security port-mode autolearn
 port-security mac-address security sticky 0002-0000-0015 vlan 1
 port-security mac-address security sticky 0002-0000-0014 vlan 1
 port-security mac-address security sticky 0002-0000-0013 vlan 1
 port-security mac-address security sticky 0002-0000-0012 vlan 1
 port-security mac-address security sticky 0002-0000-0011 vlan 1
#

```

Execute the **display port-security interface** command after the number of MAC addresses learned by the port reaches 64, and you can see that the port security mode has changed to secure. When any frame with a new MAC address arrives, intrusion protection is triggered and you can see the following trap message.

```
#Jan 14 10:39:47:135 2011 Device PORTSEC/4/VIOLATION: Trap1.3.6.1.4.1.25506.2.26.1.3.2:
```

```
An intrusion occurs!
IfIndex: 9437185
Port: 9437185
MAC Addr: 00:02:00:00:00:32
VLAN ID: 1
IfAdminStatus: 1
```

# Execute the **display interface** command, and can see that the port security feature has disabled the port.

```
[Device-GigabitEthernet1/0/1] display interface gigabitethernet 1/0/1
GigabitEthernet1/0/1 current state: DOWN (Port Security Disabled)
IP Packet Frame Type: PKTFMT_ETHNT_2, Hardware Address: 000f-cb00-5558
Description: GigabitEthernet1/0/1 Interface
.....
```

The port should be re-enabled 30 seconds later.

```
[Device-GigabitEthernet1/0/1] display interface gigabitethernet 1/0/1
GigabitEthernet1/0/1 current state: UP
IP Packet Frame Type: PKTFMT_ETHNT_2, Hardware Address: 000f-cb00-5558
Description: GigabitEthernet1/0/1 Interface
.....
```

Delete several secure MAC addresses, and you can see that the port security mode of the port changes to autoLearn, and the port can learn MAC addresses again.

## Configuring the userLoginWithOUI mode

### Network requirements

As shown in [Figure 88](#), a client is connected to the Device through port GigabitEthernet 1/0/1. The Device authenticates the client with a RADIUS server. If the authentication succeeds, the client is authorized to access the Internet.

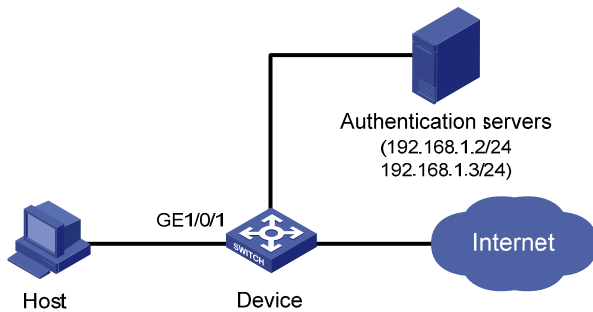
- The RADIUS server at 192.168.1.2 functions as the primary authentication server and the secondary accounting server, and the RADIUS server at 192.168.1.3 functions as the secondary authentication server and the primary accounting server. The shared key for authentication is name, and that for accounting is money.
- All users use the default authentication, authorization, and accounting methods of ISP domain **sun**, which can accommodate up to 30 users.
- The RADIUS server response timeout time is five seconds and the maximum number of RADIUS packet retransmission attempts is five. The Device sends real-time accounting packets to the RADIUS server at an interval of 15 minutes, and sends usernames without domain names to the RADIUS server.

Configure port GigabitEthernet 1/0/1 of the Device to:

- Allow only one 802.1X user to be authenticated.

- Allow up to 16 OUI values to be configured and allow one terminal that uses any of the OUI values to access the port in addition to an 802.1X user.

**Figure 88 Network diagram**



## Configuration procedure

Configurations on the host and RADIUS servers are not shown. The following configuration steps cover some AAA/RADIUS configuration commands. For more information about the commands, see *Security Command Referenced*.

1. Configure the RADIUS protocol:

# Configure a RADIUS scheme named **radsun**.

```
<Device> system-view
[Device] radius scheme radsun
[Device-radius-radsun] primary authentication 192.168.1.2
[Device-radius-radsun] primary accounting 192.168.1.3
[Device-radius-radsun] secondary authentication 192.168.1.3
[Device-radius-radsun] secondary accounting 192.168.1.2
[Device-radius-radsun] key authentication name
[Device-radius-radsun] key accounting money
[Device-radius-radsun] timer response-timeout 5
[Device-radius-radsun] retry 5
[Device-radius-radsun] timer realtime-accounting 15
[Device-radius-radsun] user-name-format without-domain
[Device-radius-radsun] quit
```

# Configure ISP domain **sun** to use RADIUS scheme **radsun** for authentication, authorization, and accounting of all types of users. Specify that the ISP domain can contain up to 30 users.

```
[Device] domain sun
[Device-isp-sun] authentication default radius-scheme radsun
[Device-isp-sun] authorization default radius-scheme radsun
[Device-isp-sun] accounting default radius-scheme radsun
[Device-isp-sun] access-limit enable 30
[Device-isp-sun] quit
```

2. Configure 802.1X:

# Set the 802.1X authentication method to CHAP. (This configuration is optional. By default, the authentication method is CHAP for 802.1X.)

```
[Device] dot1x authentication-method chap
```

3. Configure port security:

# Enable port security.

```

[Device] port-security enable
Add five OUI values.
[Device] port-security oui 1234-0100-1111 index 1
[Device] port-security oui 1234-0200-1111 index 2
[Device] port-security oui 1234-0300-1111 index 3
[Device] port-security oui 1234-0400-1111 index 4
[Device] port-security oui 1234-0500-1111 index 5
[Device] interface gigabitethernet 1/0/1
Set the port security mode to userLoginWithOUI.
[Device-GigabitEthernet1/0/1] port-security port-mode userlogin-withoui

```

## Verifying the configuration

```

Display the RADIUS scheme radsun.
<Device> display radius scheme radsun
SchemeName : radsun
Index : 1 Type : standard
Primary Auth Server:
 IP: 192.168.1.2 Port: 1812 State: active
 Encryption Key : N/A
 VPN instance : N/A
 Probe username : N/A
 Probe interval : N/A
Primary Acct Server:
 IP: 192.168.1.3 Port: 1813 State: active
 Encryption Key : N/A
 VPN instance : N/A
Second Auth Server:
 IP: 192.168.1.3 Port: 1812 State: active
 Encryption Key : N/A
 VPN instance : N/A
 Probe username : N/A
 Probe interval : N/A
Second Acct Server:
 IP: 192.168.1.2 Port: 1813 State: active
 Encryption Key : N/A
 VPN instance : N/A
Auth Server Encryption Key : *****
Acct Server Encryption Key : *****
Accounting-On packet disable, send times : 5 , interval : 3s
Interval for timeout(second) : 5
Retransmission times for timeout : 5
Interval for realtime accounting(minute) : 15
Retransmission times of realtime-accounting packet : 5
Retransmission times of stop-accounting packet : 500
Quiet-interval(min) : 5
Username format : without-domain
Data flow unit : Byte
Packet unit : one

```

# Display the configuration of the ISP domain sun.

```
<Device> display domain sun
 Domain : sun
 State : Active
 Access-limit : 30
 Accounting method : Required
 Default authentication scheme : radius:radsun
 Default authorization scheme : radius:radsun
 Default accounting scheme : radius:radsun
 Domain User Template:
 Idle-cut : Disabled
 Self-service : Disabled
 Authorization attributes:
```

# Display the port security configuration.

```
<Device> display port-security interface gigabitethernet 1/0/1
 Equipment port-security is enabled
 Trap is disabled
 Disableport Timeout: 20s
 OUI value:
 Index is 1, OUI value is 123401
 Index is 2, OUI value is 123402
 Index is 3, OUI value is 123403
 Index is 4, OUI value is 123404
 Index is 5, OUI value is 123405
```

```
GigabitEthernet1/0/1 is link-up
 Port mode is userLoginWithOUI
 NeedToKnow mode is disabled
 Intrusion Protection mode is NoAction
 Max MAC address number is not configured
 Stored MAC address number is 0
 Authorization is permitted
 Security MAC address learning mode is sticky
 Security MAC address aging type is absolute
```

After an 802.1X user gets online, you can see that the number of secure MAC addresses stored is 1.

# Display 802.1X information.

```
<Device> display dot1x interface gigabitethernet 1/0/1
 Equipment 802.1X protocol is enabled
 CHAP authentication is enabled
 EAD quick deploy is disabled

 Configuration: Transmit Period 30 s, Handshake Period 15 s
 Quiet Period 60 s, Quiet Period Timer is disabled
 Supp Timeout 30 s, Server Timeout 100 s
 Reauth Period 3600 s
 The maximal retransmitting times 2
 EAD quick deploy configuration:
```

EAD timeout: 30m

The maximum 802.1X user resource number is 1024 per slot  
Total current used 802.1X resource number is 1

GigabitEthernet1/0/1 is link-up  
802.1X protocol is enabled  
Handshake is enabled  
Handshake secure is disabled  
802.1X unicast-trigger is enabled  
Periodic reauthentication is disabled  
The port is an authenticator  
Authentication Mode is Auto  
Port Control Type is Mac-based  
802.1X Multicast-trigger is enabled  
Mandatory authentication domain: NOT configured  
Guest VLAN: NOT configured  
Auth-Fail VLAN: NOT configured  
Critical VLAN: NOT configured  
Critical recovery-action: NOT configured  
Max number of on-line users is 256

EAPOL Packet: Tx 16331, Rx 102  
Sent EAP Request/Identity Packets : 16316  
EAP Request/Challenge Packets: 6  
EAP Success Packets: 4, Fail Packets: 5  
Received EAPOL Start Packets : 6  
EAPOL LogOff Packets: 2  
EAP Response/Identity Packets : 80  
EAP Response/Challenge Packets: 6  
Error Packets: 0

1. Authenticated user : MAC address: 0002-0000-0011

Controlled User(s) amount to 1

In addition, the port allows an additional user whose MAC address has an OUI among the specified OUIs to access the port.

# Display MAC address information for interface GigabitEthernet 1/0/1.

<Device> display mac-address interface gigabitethernet 1/0/1

| MAC ADDR       | VLAN ID | STATE   | PORT INDEX           | AGING TIME(s) |
|----------------|---------|---------|----------------------|---------------|
| 1234-0300-0011 | 1       | Learned | GigabitEthernet1/0/1 | AGING         |

--- 1 mac address(es) found ---

# Configuring the macAddressElseUserLoginSecure mode

## Network requirements

As shown in [Figure 88](#), a client is connected to the Device through GigabitEthernet 1/0/1. The Device authenticates the client by a RADIUS server. If the authentication succeeds, the client is authorized to access the Internet.

Restrict port GigabitEthernet 1/0/1 of the Device:

- Allow more than one MAC authenticated user to log on.
- For 802.1X users, perform MAC authentication first and then, if MAC authentication fails, 802.1X authentication. Allow only one 802.1X user to log on.
- Set fixed username and password for MAC authentication. Set the total number of MAC authenticated users and 802.1X authenticated users to 64.
- Enable NTK to prevent frames from being sent to unknown MAC addresses.

## Configuration procedure

Configurations on the host and RADIUS servers are not shown.

1. Configure the RADIUS protocol:

Configure the RADIUS authentication/accounting and ISP domain settings the same as in [Configuring the userLoginWithOUI mode](#).

2. Configure port security:

# Enable port security.

```
<Device> system-view
[Device] port-security enable
```

# Configure a MAC authentication user, setting the username and password to aaa and 123456 respectively.

```
[Device] mac-authentication user-name-format fixed account aaa password simple 123456
[Device] interface gigabitethernet 1/0/1
```

# Specify ISP domain **sun** for MAC authentication.

```
[Device] mac-authentication domain sun
[Device] interface gigabitethernet 1/0/1
```

# Set the 802.1X authentication method to CHAP. (This configuration is optional. By default, the authentication method is CHAP for 802.1X.)

```
[Device] dot1x authentication-method chap
```

# Set port security's limit on the number of MAC addresses to 64 on the port.

```
[Device-GigabitEthernet1/0/1] port-security max-mac-count 64
```

# Set the port security mode to macAddressElseUserLoginSecure.

```
[Device-GigabitEthernet1/0/1] port-security port-mode mac-else-userlogin-secure
```

# Set the NTK mode of the port to ntkonly.

```
[Device-GigabitEthernet1/0/1] port-security ntk-mode ntkonly
```

## Verifying the configuration

# Display the port security configuration.

```
<Device> display port-security interface gigabitethernet 1/0/1
Equipment port-security is enabled
Trap is disabled
```

Disableport Timeout: 20s

OUI value:

GigabitEthernet1/0/1 is link-up

Port mode is macAddressElseUserLoginSecure

NeedToKnow mode is NeedToKnowOnly

Intrusion Protection mode is NoAction

Max MAC address number is 64

Stored MAC address number is 0

Authorization is permitted

Security MAC address learning mode is sticky

Security MAC address aging type is absolute

### # Display MAC authentication information.

<Device> display mac-authentication interface gigabitethernet 1/0/1

MAC address authentication is enabled.

User name format is fixed account

Fixed username:aaa

Fixed password: \*\*\*\*\*

Offline detect period is 60s

Quiet period is 5s

Server response timeout value is 100s

The max allowed user number is 1024 per slot

Current user number amounts to 3

Current domain is mac

Silent MAC User info:

| MAC Addr | From Port | Port Index |
|----------|-----------|------------|
|----------|-----------|------------|

GigabitEthernet1/0/1 is link-up

MAC address authentication is enabled

Authenticate success: 3, failed: 7

Max number of on-line users is 256

Current online user number is 3

| MAC ADDR       | Authenticate state        | Auth Index |
|----------------|---------------------------|------------|
| 1234-0300-0011 | MAC_AUTHENTICATOR_SUCCESS | 13         |
| 1234-0300-0012 | MAC_AUTHENTICATOR_SUCCESS | 14         |
| 1234-0300-0013 | MAC_AUTHENTICATOR_SUCCESS | 15         |

### # Display 802.1X authentication information.

<Device> display dot1x interface gigabitethernet 1/0/1

Equipment 802.1X protocol is enabled

CHAP authentication is enabled

EAD quick deploy is disabled

|                                |       |                                |       |
|--------------------------------|-------|--------------------------------|-------|
| Configuration: Transmit Period | 30 s, | Handshake Period               | 15 s  |
| Quiet Period                   | 60 s, | Quiet Period Timer is disabled |       |
| Supp Timeout                   | 30 s, | Server Timeout                 | 100 s |



```
The maximal retransmitting times 2
EAD quick deploy configuration:
 EAD timeout: 30m

Total maximum 802.1X user resource number is 1024 per slot
Total current used 802.1X resource number is 1
```

```
GigabitEthernet1/0/1 is link-up
 802.1X protocol is enabled
 Handshake is enabled
 Handshake secure is disabled
 802.1X unicast-trigger is enabled
 Periodic reauthentication is disabled
 The port is an authenticator
 Authentication Mode is Auto
 Port Control Type is Mac-based
 802.1X Multicast-trigger is enabled
 Mandatory authentication domain: NOT configured
 Guest VLAN: NOT configured
 Auth-Fail VLAN: NOT configured
 Critical VLAN: NOT configured
 Critical recovery-action: NOT configured
 Max number of on-line users is 256
```

```
EAPOL Packet: Tx 16331, Rx 102
Sent EAP Request/Identity Packets : 16316
 EAP Request/Challenge Packets: 6
 EAP Success Packets: 4, Fail Packets: 5
Received EAPOL Start Packets : 6
 EAPOL LogOff Packets: 2
 EAP Response/Identity Packets : 80
 EAP Response/Challenge Packets: 6
 Error Packets: 0
1. Authenticated user : MAC address: 0002-0000-0011
```

Controlled User(s) amount to 1

As NTK is enabled, frames with unknown destination MAC addresses, multicast addresses, and broadcast addresses will be discarded.

## Troubleshooting port security

### Cannot set the port security mode

#### Symptom

Cannot set the port security mode.

```
[Device-GigabitEthernet1/0/1] port-security port-mode autolearn
```

Error:When we change port-mode, we should first change it to noRestrictions, then change it to the other.

## Analysis

For a port operating in a port security mode other than noRestrictions, you cannot change the port security mode by using the **port-security port-mode** command directly.

## Solution

Set the port security mode to noRestrictions first.

```
[Device-GigabitEthernet1/0/1] undo port-security port-mode
[Device-GigabitEthernet1/0/1] port-security port-mode autolearn
```

# Cannot configure secure MAC addresses

## Symptom

Cannot configure secure MAC addresses.

```
[Device-GigabitEthernet1/0/1] port-security mac-address security 1-1-2 vlan 1
Error: Security MAC address configuration failed.
```

## Analysis

No secure MAC address can be configured on a port operating in a port security mode other than autoLearn.

## Solution

Set the port security mode to autoLearn.

```
[Device-GigabitEthernet1/0/1] undo port-security port-mode
[Device-GigabitEthernet1/0/1] port-security max-mac-count 64
[Device-GigabitEthernet1/0/1] port-security port-mode autolearn
[Device-GigabitEthernet1/0/1] port-security mac-address security 1-1-2 vlan 1
```

# Cannot change port security mode when a user is online

## Symptom

Port security mode cannot be changed when an 802.1X authenticated or MAC authenticated user is online.

```
[DeviceGigabitEthernet1/0/1] undo port-security port-mode
Error:Cannot configure port-security for there is 802.1X user(s) on line on port
GigabitEthernet1/0/1.
```

## Analysis

Changing port security mode is not allowed when an 802.1X authenticated or MAC authenticated user is online.

## Solution

Use the **cut** command to forcibly disconnect the user from the port before changing the port security mode.

```
[Device-GigabitEthernet1/0/1] quit
[Device] cut connection interface gigabitethernet 1/0/1
[Device] interface gigabitethernet 1/0/1
```

```
[Device-GigabitEthernet1/0/1] undo port-security port-mode
```

---

# Configuring a user profile

## User profile overview

A user profile provides a configuration template to save predefined configurations, such as a Quality of Service (QoS) policy. Different user profiles are applicable to different application scenarios.

The user profile supports working with 802.1X authentication and portal authentication. It is capable of restricting authenticated users' behaviors. After the authentication server verifies a user, it sends the device the name of the user profile that is associated with the user. Then the device applies the configurations in the user profile if the profile is enabled, and allows user access based on all valid configurations. If the user profile is not enabled, the device denies the user access. After the user logs out, the device automatically disables the configurations in the user profile, and the restrictions on the users are removed.

Without user profiles, service applications are based on interface, VLAN, or globally, and a policy applies to any user that accesses the interface, or VLAN, or device. If a user moves between ports to access a device, to restrict the user behavior, you must remove the policy from the previous port and then configure the same policy on the port that the user uses. The configuration task is tedious and error prone.

User profiles provide flexible user-based service applications because a user profile is associated with a target user. Every time the user accesses the device, the device automatically applies the configurations in the associated user profile.

## User profile configuration task list

| Task                                    | Remarks  |
|-----------------------------------------|----------|
| <a href="#">Creating a user profile</a> | Required |
| <a href="#">Applying a QoS policy</a>   | Required |
| <a href="#">Enabling a user profile</a> | Required |

## Creating a user profile

### Configuration prerequisites

Before you create a user profile, complete the following tasks:

- Configure authentication parameters on the device.
- Perform configurations on the client, the device, and the authentication server, for example, username, password, authentication scheme, domain, and binding a user profile with a user.

### Configuration procedure

| Step                                          | Command                                 | Remarks                                                                |
|-----------------------------------------------|-----------------------------------------|------------------------------------------------------------------------|
| 1. Enter system view.                         | <b>system-view</b>                      | N/A                                                                    |
| 2. Create a user profile, and enter its view. | <b>user-profile</b> <i>profile-name</i> | You can use the command to enter the view of an existing user profile. |

## Applying a QoS policy

You can apply QoS policies in user profile view to implement traffic management functions.

### Configuration guidelines

- After a user profile is created, apply a QoS policy in user profile view to implement restrictions on online users. The QoS policy takes effect when the user profile is enabled and a user using the user profile goes online.
- The QoS policies that can be applied to user profiles support only the **remark**, **car**, and **filter** actions.
- Do not apply an empty policy in user profile view because a user profile with an empty policy applied cannot be enabled.
- If a user profile is enabled, you cannot modify the applied QoS policy (including the ACL that is referenced by the QoS policy) or remove it.
- For information about QoS policy configurations, see *ACL and QoS Configuration Guide*.

### Configuration procedure

| Step                        | Command                                                                            | Remarks                                                                                                                                                                                                                                                                                                                          |
|-----------------------------|------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.       | <b>system-view</b>                                                                 | N/A                                                                                                                                                                                                                                                                                                                              |
| 2. Enter user profile view. | <b>user-profile</b> <i>profile-name</i>                                            | N/A                                                                                                                                                                                                                                                                                                                              |
| 3. Apply the QoS policy.    | <b>qos apply policy</b> <i>policy-name</i><br>{ <b>inbound</b>   <b>outbound</b> } | The <b>inbound</b> keyword applies the QoS policy to incoming traffic of the switch (traffic sent by online users).<br>The <b>outbound</b> keyword applies the QoS policy to outgoing traffic of the switch (traffic sent to online users).<br><br>The <b>outbound</b> keyword is not available on the HP 5500 SI Switch Series. |

## Enabling a user profile

Enable a user profile so that configurations in the profile can be applied by the device to restrict user behaviors. If the device detects that the user profile is disabled, the device denies the associated user even the user has been verified by the authentication server.

To enable a user profile:

| Step                      | Command                                               | Remarks                                |
|---------------------------|-------------------------------------------------------|----------------------------------------|
| 1. Enter system view.     | <b>system-view</b>                                    | N/A                                    |
| 2. Enable a user profile. | <b>user-profile</b> <i>profile-name</i> <b>enable</b> | A user profile is disabled by default. |

**NOTE:**

- You can only edit or remove the configurations in a disabled user profile.
- Disabling a user profile logs out the users that are using the user profile.

## Displaying and maintaining user profiles

| Task                                                     | Command                                                                                                        | Remarks               |
|----------------------------------------------------------|----------------------------------------------------------------------------------------------------------------|-----------------------|
| Display information about all the created user profiles. | <b>display user-profile</b> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] | Available in any view |

---

# Configuring password control

## Password control overview

Password control refers to a set of functions provided by the local authentication server to control user login passwords, super passwords, and user login status based on predefined policies. The rest of this section describes the password control functions in detail.

1. Minimum password length

By setting a minimum password length, you can enforce users to use passwords long enough for system security. If a user specifies a shorter password, the system rejects the setting and prompts the user to re-specify a password.

2. Minimum password update interval

This function allows you to set the minimum interval at which users can change their passwords. If a non-manage level user logs in to change the password but the time that elapses since the last change is less than this interval, the system denies the request. For example, if you set this interval to 48 hours, a non-manage level user cannot change the password twice within 48 hours. This prevents users from changing their passwords frequently.

---

**NOTE:**

- This function is not effective for users of the manage level. For information about user levels, see *Fundamentals Configuration Guide*.
- This function is not effective for a user who is prompted to change the password at the first login or a user whose password has just been aged out.

---

3. Password aging

Password aging imposes a lifecycle on a user password. After the password aging time expires, the user needs to change the password.

If a user enters an expired password when logging in, the system displays an error message and prompts the user to provide a new password and to confirm it by entering it again. The new password must be a valid one and the user must enter exactly the same password when confirming it.

4. Early notice on pending password expiration

When a user logs in, the system checks whether the password will expire in a time equal to or less than the specified period. If so, the system notifies the user of the expiry time and provides a choice for the user to change the password. If the user provides a new password that is qualified, the system records the new password and the time. If the user chooses to leave the password or the user fails to change it, the system allows the user to log in using the present password.

---

**NOTE:**

Telnet, SSH, and terminal users can change their passwords by themselves, while FTP users can only have their passwords changed by the administrator.

---

5. Login with an expired password

You can allow a user to log in a certain number of times within a specified period of time after the password expires, so that the user does not need to change the password immediately. For

example, if you set the maximum number of logins with an expired password to three and the time period to 15 days, a user can log in three times within 15 days after the password expires.

## 6. Password history

With this feature enabled, the system maintains certain entries of passwords that a user has used. When a user changes the password, the system checks the new password against the used ones. The new password must be different from the used ones by at least four characters and the four characters must not be the same. Otherwise, you will fail to change the password and the system displays an error message.

You can set the maximum number of history password records for the system to maintain for each user. When the number of history password records exceeds your setting, the latest record will overwrite the earliest one.

## 7. Login attempt limit

Limiting the number of consecutive failed login attempts can effectively prevent password guessing.

If an FTP or virtual terminal line (VTY) user fails authentication due to a password error, the system adds the user to a password control blacklist. If a user fails to provide the correct password after the specified number of consecutive attempts, the system takes action as configured:

- Prohibiting the user from logging in until the user is removed from the password control blacklist manually.
- Allowing the user to try continuously and removing the user from the password control blacklist when the user logs in to the system successfully or the blacklist entry times out (the blacklist entry aging time is one minute).
- Prohibiting the user from logging in within a configurable period of time, and allowing the user to log in again after the period of time elapses or the user is removed from the password control blacklist.

A password control blacklist can contain up to 1024 entries.

A login attempt using a wrong username will undoubtedly fail but the username will not be added to the password control blacklist.

Web users failing login authentication are not added to the password control blacklist. Users accessing the system through the console or AUX interface are not blacklisted either, because the system is unable to obtain the IP addresses of these users and these users are privileged and therefore relatively secure to the system.

## 8. Password composition checking

A password can be a combination of characters from the following four categories:

- Uppercase letters A to Z
- Lowercase letters a to z
- Digits 0 to 9
- 32 special characters including blank space and ~`!@#\$\$%^&\*()\_+={}|[]\:"';<>,. /.

Depending on the system security requirements, you can set the minimum number of categories a password must contain and the minimum number of characters of each category.

There are four password combination levels: 1, 2, 3, and 4, each representing the number of categories that a password must at least contain. Level 1 means that a password must contain characters of one category, level 2 at least two categories, and so on.

When a user sets or changes the password, the system checks if the password satisfies the composition requirement. If not, the system displays an error message.



## 9. Password complexity checking

A less complicated password such as a password containing the username or repeated characters is more likely to be cracked. For higher security, you can configure a password complexity checking policy to make sure that all user passwords are relatively complicated. With such a policy configured, when a user configures a password, the system checks the complexity of the password. If the password is not qualified, the system refuses the password and displays a password configuration failure message.

You can impose the following password complexity requirements:

- A password cannot contain the username or the reverse of the username. For example, if the username is abc, a password such as abc982 or 2cba is unqualified.
- No character of the password is repeated three or more times consecutively. For example, password a111 is not qualified.

## 10. Password display in the form of a string of \*

For the sake of security, the password a user enters is displayed in the form of a string of \*.

## 11. Authentication timeout management

The authentication period is from when the server obtains the username to when the server finishes authenticating the user's password. If a Telnet user fails to log in within the configured period of time, the system tears down the connection.

## 12. Maximum account idle time

You can set the maximum account idle time to make accounts staying idle for this period of time become invalid and unable to log in again. For example, if you set the maximum account idle time to 60 days and user using the account **test** has never logged in successfully within 60 days after the last successful login, the account becomes invalid.

## 13. Logging

The system logs all successful password changing events and the events of adding users to the password control blacklist.

# Password control configuration task list

The password control functions can be configured in several views, and different views support different functions. The settings configured in different views or for different objects have different application ranges and different priorities:

- Global settings in system view apply to all local user passwords and super passwords.
- Settings in user group view apply to the passwords of all local users in the user group.
- Settings in local user view apply to only the password of the local user.
- Settings for super passwords apply to only super passwords.

The above four types of settings have different priorities:

- For local user passwords, the settings with a smaller application range have a higher priority.
- For super passwords, the settings configured specifically for super passwords, if any, override those configured in system view.

Complete the following tasks to configure password control:

| Task                      | Remarks  |
|---------------------------|----------|
| Enabling password control | Required |

| Task                                              | Remarks  |
|---------------------------------------------------|----------|
| Setting global password control parameters        | Optional |
| Setting user group password control parameters    | Optional |
| Setting local user password control parameters    | Optional |
| Setting super password control parameters         | Optional |
| Setting a local user password in interactive mode | Optional |

## Configuring password control

### Enabling password control

To enable password control functions, you need to:

1. Enable the password control feature in system view. Only after the password control feature is enabled globally, can password control configurations take effect.
2. Enable password control functions. Some password control functions need to be enabled individually after the password control feature is enabled globally. These functions include:
  - o Password aging
  - o Minimum password length
  - o Password history
  - o Password composition checking

You must enable a function for its relevant configurations to take effect.

To enable password control:

| Step                                                | Command                                                                   | Remarks                                                                        |
|-----------------------------------------------------|---------------------------------------------------------------------------|--------------------------------------------------------------------------------|
| 1. Enter system view.                               | <b>system-view</b>                                                        | N/A                                                                            |
| 2. Enable the password control feature.             | <b>password-control enable</b>                                            | Disabled by default                                                            |
| 3. Enable a password control function individually. | <b>password-control { aging   composition   history   length } enable</b> | Optional<br>All of the four password control functions are enabled by default. |

#### NOTE:

After global password control is enabled, local user passwords configured on the device are not displayed when you use the corresponding display command.

### Setting global password control parameters

| Step                  | Command            | Remarks |
|-----------------------|--------------------|---------|
| 1. Enter system view. | <b>system-view</b> | N/A     |

| Step                                                                                                                                           | Command                                                                                                                                         | Remarks                                                                                                                                                                             |
|------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2. Set the password aging time.                                                                                                                | <b>password-control aging</b> <i>aging-time</i>                                                                                                 | Optional<br>90 days by default                                                                                                                                                      |
| 3. Set the minimum password update interval.                                                                                                   | <b>password-control password update interval</b> <i>interval</i>                                                                                | Optional<br>24 hours by default                                                                                                                                                     |
| 4. Set the minimum password length.                                                                                                            | <b>password-control length</b> <i>length</i>                                                                                                    | Optional<br>10 characters by default                                                                                                                                                |
| 5. Configure the password composition policy.                                                                                                  | <b>password-control composition type-number</b> <i>policy-type</i><br>[ <b>type-length</b> <i>type-length</i> ]                                 | Optional<br>By default, the minimum number of password composition types is 1 and the minimum number of characters of a password composition type is 1 too.                         |
| 6. Configure the password complexity checking policy.                                                                                          | <b>password-control complexity</b><br>{ <b>same-character</b>   <b>user-name</b> }<br><b>check</b>                                              | Optional<br>By default, the system does not perform password complexity checking.                                                                                                   |
| 7. Set the maximum number of history password records for each user.                                                                           | <b>password-control history</b><br><i>max-record-num</i>                                                                                        | Optional<br>4 by default                                                                                                                                                            |
| 8. Specify the maximum number of login attempts and the action to be taken when a user fails to log in after the specified number of attempts. | <b>password-control login-attempt</b><br><i>login-times</i> [ <b>exceed</b> { <b>lock</b>   <b>unlock</b> }<br>  <b>lock-time</b> <i>time</i> ] | Optional<br>By default, the maximum number of login attempts is 3 and a user failing to log in after the specified number of attempts must wait for one minute before trying again. |
| 9. Set the number of days during which the user is warned of the pending password expiration.                                                  | <b>password-control alert-before-expire</b> <i>alert-time</i>                                                                                   | Optional<br>7 days by default                                                                                                                                                       |
| 10. Set the maximum number of days and maximum number of times that a user can log in after the password expires.                              | <b>password-control expired-user-login delay</b> <i>delay</i><br><b>times</b> <i>times</i>                                                      | Optional<br>By default, a user can log in three times within 30 days after the password expires.                                                                                    |
| 11. Set the authentication timeout time.                                                                                                       | <b>password-control authentication-timeout</b><br><i>authentication-timeout</i>                                                                 | Optional<br>60 seconds by default                                                                                                                                                   |
| 12. Set the maximum account idle time.                                                                                                         | <b>password-control login idle-time</b><br><i>idle-time</i>                                                                                     | Optional<br>90 days by default                                                                                                                                                      |

**NOTE:**

The specified action to be taken after a user fails to log in for the specified number of attempts takes effect immediately, and can thus affect the users already in the password control blacklist. Other password control configurations take effect only for users logging in later and passwords configured later.

## Setting user group password control parameters

| Step                                                             | Command                                                                                                                   | Remarks                                                                                    |
|------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|
| 1. Enter system view.                                            | <b>system-view</b>                                                                                                        | N/A                                                                                        |
| 2. Create a user group and enter user group view.                | <b>user-group</b> <i>group-name</i>                                                                                       | N/A                                                                                        |
| 3. Configure the password aging time for the user group.         | <b>password-control aging</b> <i>aging-time</i>                                                                           | Optional<br>By default, the password aging time configured in system view is used.         |
| 4. Configure the minimum password length for the user group.     | <b>password-control length</b> <i>length</i>                                                                              | Optional<br>By default, the minimum password length configured in system view is used.     |
| 5. Configure the password composition policy for the user group. | <b>password-control composition</b><br><b>type-number</b> <i>type-number</i><br>[ <b>type-length</b> <i>type-length</i> ] | Optional<br>By default, the password composition policy configured in system view is used. |

## Setting local user password control parameters

| Step                                                             | Command                                                                                                                   | Remarks                                                                                                                                                                                                     |
|------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                                            | <b>system-view</b>                                                                                                        | N/A                                                                                                                                                                                                         |
| 2. Create a local user and enter local user view.                | <b>local-user</b> <i>user-name</i>                                                                                        | N/A                                                                                                                                                                                                         |
| 3. Configure the password aging time for the local user.         | <b>password-control aging</b> <i>aging-time</i>                                                                           | Optional<br>By default, the setting for the user group to which the local user belongs is used; if no aging time is configured for the user group, the setting in system view is used.                      |
| 4. Configure the minimum password length for the local user.     | <b>password-control length</b> <i>length</i>                                                                              | Optional<br>By default, the setting for the user group to which the local user belongs is used; if no minimum password length is configured for the user group, the setting in system view is used.         |
| 5. Configure the password composition policy for the local user. | <b>password-control composition</b><br><b>type-number</b> <i>type-number</i><br>[ <b>type-length</b> <i>type-length</i> ] | Optional<br>By default, the settings for the user group to which the local user belongs are used; if no password composition policy is configured for the user group, the settings in system view are used. |

## Setting super password control parameters

CLI commands fall into four levels: visit, monitor, system, and manage, in ascending order. Accordingly, login users fall into four levels, each corresponding to a command level. A user of a certain level can only use the commands at that level or lower levels.

To switch from a lower user level to a higher one, a user needs to enter a password for authentication. This password is called a super password. For more information on super passwords, see *Fundamentals Configuration Guide*.

To set super password control parameters:

| Step                                                              | Command                                                                                                           | Remarks                                                                                                                                                     |
|-------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                                             | <b>system-view</b>                                                                                                | N/A                                                                                                                                                         |
| 2. Set the password aging time for super passwords.               | <b>password-control super aging</b><br><i>aging-time</i>                                                          | Optional<br>90 days by default                                                                                                                              |
| 3. Configure the minimum length for super passwords.              | <b>password-control super length</b><br><i>length</i>                                                             | Optional<br>10 characters by default                                                                                                                        |
| 4. Configure the password composition policy for super passwords. | <b>password-control super composition type-number</b><br><i>type-number [ type-length</i><br><i>type-length ]</i> | Optional<br>By default, the minimum number of password composition types is 1 and the minimum number of characters of a password composition type is 1 too. |

## Setting a local user password in interactive mode

You can set a password for a local user in interactive mode. When doing so, you need to confirm the password.

To set a password for a local user in interactive mode:

| Step                                                        | Command                            |
|-------------------------------------------------------------|------------------------------------|
| 1. Enter system view.                                       | <b>system-view</b>                 |
| 2. Create a local user and enter local user view.           | <b>local-user</b> <i>user-name</i> |
| 3. Set the password for the local user in interactive mode. | <b>password</b>                    |

## Displaying and maintaining password control

| Task                                                | Command                                                                                                                            | Remarks               |
|-----------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Display password control configuration information. | <b>display password-control [ super ]</b><br>[   { <b>begin</b>   <b>exclude</b>   <b>include</b> }<br><i>regular-expression</i> ] | Available in any view |

| Task                                                               | Command                                                                                                                                                                                                                                     | Remarks                |
|--------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|
| Display information about users in the password control blacklist. | <b>display password-control blacklist</b><br>[ <b>user-name</b> <i>name</i>   <b>ip</b><br><i>ipv4-address</i>   <b>ipv6</b> <i>ipv6-address</i> ]<br>[   { <b>begin</b>   <b>exclude</b>   <b>include</b> }<br><i>regular-expression</i> ] | Available in any view  |
| Delete users from the password control blacklist.                  | <b>reset password-control blacklist</b><br>[ <b>user-name</b> <i>name</i> ]                                                                                                                                                                 | Available in user view |
| Clear history password records.                                    | <b>reset password-control</b><br><b>history-record</b> [ <b>user-name</b> <i>name</i>  <br><b>super</b> [ <b>level</b> <i>level</i> ] ]                                                                                                     | Available in user view |

#### NOTE:

The **reset password-control history-record** command can delete the history password records of a specific user or all users even when the password history function is disabled.

## Password control configuration example

### Network requirements

Implementing the following global password control policy:

- An FTP or VTY user failing to provide the correct password in two successive login attempts is permanently prohibited from logging in.
- A user can log in five times within 60 days after the password expires.
- The password aging time is 30 days.
- The minimum password update interval is 36 hours.
- The maximum account idle time is 30 days.
- A password cannot contain the username or the reverse of the username.
- No character occurs consecutively three or more times in a password.

Implementing the following super password control policy:

- A super password must contain at least three types of valid characters, five or more of each type.

Implementing the following password control policy for local Telnet user **test**:

- The password must contain at least 12 characters.
- The password must consist of at least two types of valid characters, five or more of each type.
- The password aging time is 20 days.

### Configuration procedure

# Enable the password control feature globally.

```
<Sysname> system-view
[Sysname] password-control enable
```

# Prohibit the user from logging in forever after two successive login failures.

```
[Sysname] password-control login-attempt 2 exceed lock
```

# Set the password aging time to 30 days for all passwords.

```
[Sysname] password-control aging 30
```

```

Set the minimum password update interval to 36 hours.
[Sysname] password-control password update interval 36

Specify that a user can log in five times within 60 days after the password expires.
[Sysname] password-control expired-user-login delay 60 times 5

Set the maximum account idle time to 30 days.
[Sysname] password-control login idle-time 30

Refuse any password that contains the username or the reverse of the username.
[Sysname] password-control complexity user-name check

Specify that no character of the password can be repeated three or more times consecutively.
[Sysname] password-control complexity same-character check

Set the minimum number of composition types for super passwords to 3 and the minimum number of
characters of each composition type to 5.
[Sysname] password-control super composition type-number 3 type-length 5

Configure a super password.
[Sysname] super password level 3 simple 12345ABGFTweuix

Create a local user named test.
[Sysname] local-user test

Set the service type of the user to Telnet.
[Sysname-user-test] service-type telnet

Set the minimum password length to 12 for the local user.
[Sysname-user-test] password-control length 12

Set the minimum number of password composition types to 2 and the minimum number of characters
of each password composition type to 5 for the local user.
[Sysname-user-test] password-control composition type-number 2 type-length 5

Set the password aging time to 20 days for the local user.
[Sysname-user-test] password-control aging 20

Configure the password of the local user in interactive mode.
[Sysname-user-test] password
Password:*****
Confirm :*****
Updating user(s) information, please wait.....
[Sysname-user-test] quit

```

## Verifying the configuration

```

Display the global password control configuration information.
<Sysname> display password-control
Global password control configurations:
 Password control: Enabled
 Password aging: Enabled (30 days)
 Password length: Enabled (10 characters)
 Password composition: Enabled (1 types, 1 characters per type)
 Password history: Enabled (max history record:4)
 Early notice on password expiration: 7 days
 User authentication timeout: 60 seconds

```

```
Maximum failed login attempts: 2 times
Login attempt-failed action: Lock
Minimum password update time: 36 hours
User account idle-time: 30 days
Login with aged password: 5 times in 60 day(s)
Password complexity: Enabled (username checking)
 Enabled (repeated characters checking)
```

# Display the password control configuration information for super passwords.

```
<Sysname> display password-control super
Super password control configurations:
Password aging: Enabled (30 days)
Password length: Enabled (10 characters)
Password composition: Enabled (3 types, 5 characters per type)
```

# Display the password control configuration information for local user **test**.

```
<Sysname> display local-user user-name test
The contents of local user test:
State: Active
ServiceType: telnet
Access-limit: Disable Current AccessNum: 0
User-group: system
Bind attributes:
Authorization attributes:
Password aging: Enabled (20 days)
Password length: Enabled (12 characters)
Password composition: Enabled (2 types, 5 characters per type)
Total 1 local user(s) matched.
```



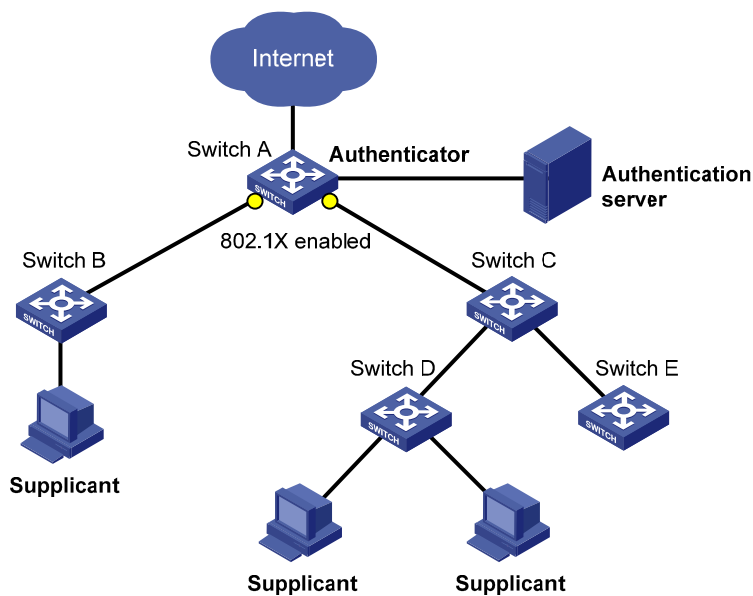
# Configuring HABP

## HABP overview

The HW Authentication Bypass Protocol (HABP) is intended to enable the downstream network devices of an access device to bypass 802.1X authentication and MAC authentication configured on the access device.

As shown in Figure 89, 802.1X authenticator Switch A has two switches attached to it: Switch B and Switch C. On Switch A, 802.1X authentication is enabled globally and on the ports connecting the downstream network devices. The end-user devices (the supplicants) run the 802.1X client software for 802.1X authentication. For Switch B and Switch D, where the 802.1X client is not supported (which is typical of network devices), the communication between them will fail because they cannot pass 802.1X authentication and their packets will be blocked on Switch A. To allow the two switches to communicate, you can use HABP.

Figure 89 Network diagram for HABP application



HABP is a link layer protocol that works above the MAC layer. It is built on the client-server model. Generally, the HABP server is enabled on the authentication device (which is configured with 802.1X or MAC authentication, such as Switch A in the above example), and the attached switches function as the HABP clients, such as Switch B through Switch E in the example. No device can function as both an HABP server and a client at the same time. Typically, the HABP server sends HABP requests to all its clients periodically to collect their MAC addresses, and the clients respond to the requests. After the server learns the MAC addresses of all the clients, it registers the MAC addresses as HABP entries. Then, link layer frames exchanged between the clients can bypass the 802.1X authentication on ports of the server without affecting the normal operation of the whole network. All HABP packets must travel in a specified VLAN. Communication between the HABP server and HABP clients is implemented through this VLAN.

In a cluster, if a member switch with 802.1X authentication or MAC authentication enabled is attached with some other member switches of the cluster, you also need to configure HABP server on this device.

Otherwise, the cluster management device will not be able to manage the devices attached to this member switch. For more information about the cluster function, see *Network Management and Monitoring Configuration Guide*.

## Configuring HABP

### Configuring the HABP server

An HABP server is usually configured on the authentication device enabled with 802.1X authentication or MAC address authentication. The HABP server sends HABP requests to the attached switches (HABP clients) at a specified interval, collecting their MAC addresses from the responses. HABP packets are transmitted in the VLAN specified on the HABP server.

To configure an HABP server:

| Step                                                                            | Command                                | Remarks                                                                                                                                                                 |
|---------------------------------------------------------------------------------|----------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                                                           | <b>system-view</b>                     | N/A                                                                                                                                                                     |
| 2. Enable HABP.                                                                 | <b>habp enable</b>                     | Optional<br>Enabled by default                                                                                                                                          |
| 3. Configure HABP to work in server mode and specify the VLAN for HABP packets. | <b>habp server vlan</b> <i>vlan-id</i> | HABP works in client mode by default.<br>The VLAN specified on the HABP server for transmitting HABP packets must be the same as that to which the HABP clients belong. |
| 4. Set the interval to send HABP requests.                                      | <b>habp timer</b> <i>interval</i>      | Optional<br>20 seconds by default                                                                                                                                       |

### Configuring an HABP client

An HABP client is usually configured on each device that is attached to the authentication device. After receiving an HABP request from the HABP server, an HABP client responds to the request, delivering its MAC address to the server, and forwards the HABP request to its attached switches. HABP packets are transmitted in the VLAN to which the HABP client belongs.

To configure an HABP client:

| Step                                      | Command                 | Remarks                                           |
|-------------------------------------------|-------------------------|---------------------------------------------------|
| 1. Enter system view.                     | <b>system-view</b>      | N/A                                               |
| 2. Enable HABP.                           | <b>habp enable</b>      | Optional<br>Enabled by default                    |
| 3. Configure HABP to work in client mode. | <b>undo habp server</b> | Optional<br>HABP works in client mode by default. |

| Step                                                  | Command                                | Remarks                                                                                                                                                                                    |
|-------------------------------------------------------|----------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 4. Specify the VLAN to which the HABP client belongs. | <b>habp client vlan</b> <i>vlan-id</i> | Optional<br>By default, an HABP client belongs to VLAN 1.<br>The VLAN to which an HABP client belongs must be the same as that specified on the HABP server for transmitting HABP packets. |

## Displaying and maintaining HABP

| Task                                    | Command                                                                                                        | Remarks               |
|-----------------------------------------|----------------------------------------------------------------------------------------------------------------|-----------------------|
| Display HABP configuration information. | <b>display habp</b> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]         | Available in any view |
| Display HABP MAC address table entries. | <b>display habp table</b> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]   | Available in any view |
| Display HABP packet statistics.         | <b>display habp traffic</b> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] | Available in any view |

## HABP configuration example

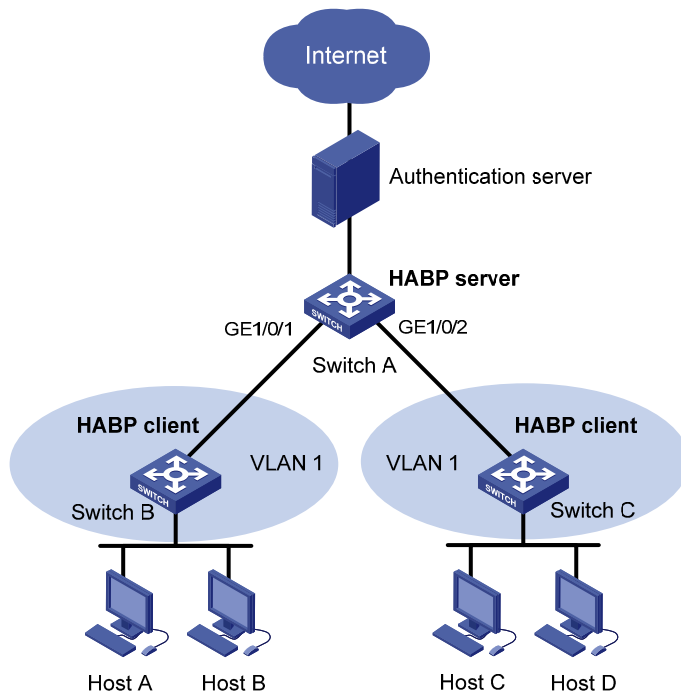
### Network requirements

As shown in [Figure 90](#), Switch A is attached with access devices Switch B and Switch C. 802.1X authentication is configured on Switch A for central authentication and management of users (Host A through Host D).

For communication between Switch B and Switch C, enable HABP server on Switch A, enable HABP client on Switch B and Switch C, and specify VLAN 1 for HABP packets.

Configure the HABP server to send HABP request packets to the HABP clients in VLAN 1 at an interval of 50 seconds.

Figure 90 Network diagram



## Configuration procedure

1. Configure Switch A:
  - # Perform 802.1X related configurations on Switch A (see "[Configuring 802.1X](#)").
  - # Enable HABP. (HABP is enabled by default. This configuration is optional.)

```
<SwitchA> system-view
[SwitchA] habp enable
```
  - # Configure HABP to work in server mode, and specify VLAN 1 for HABP packets.

```
[SwitchA] habp server vlan 1
```
  - # Set the interval at which the switch sends HABP request packets to 50 seconds.

```
[SwitchA] habp timer 50
```
2. Configure Switch B:
  - # Enable HABP. (HABP is enabled by default. This configuration is optional.)

```
<SwitchA> system-view
[SwitchB] habp enable
```
  - # Configure HABP to work in client mode. (HABP works in client mode by default. This configuration is optional.)

```
[SwitchB] undo habp server
```
  - # Specify the VLAN to which the HABP client belongs as VLAN 1. (An HABP client belongs to VLAN 1 by default. This configuration is optional.)

```
[SwitchB] habp client vlan 1
```
3. Configure Switch C:

Configurations on Switch C are similar to those on Switch B.
4. Verify your configuration:
  - # Display HABP configuration information.

```
<SwitchA> display habp
Global HABP information:
 HABP Mode: Server
 Sending HABP request packets every 50 seconds
 Bypass VLAN: 1
```

# Display HABP MAC address table entries.

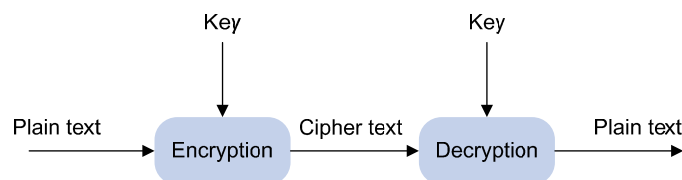
```
<SwitchA> display habp table
MAC Holdtime Receive Port
001f-3c00-0030 53 GigabitEthernet1/0/2
001f-3c00-0031 53 GigabitEthernet1/0/1
```

# Managing public keys

## Overview

To protect data confidentiality during transmission, the data sender uses an algorithm and a key to encrypt the plain text data before sending the data out, and the receiver uses the same algorithm with the help of a key to decrypt the data, as shown in [Figure 91](#).

**Figure 91 Encryption and decryption**



The keys that participate in the conversion between the plain text and the cipher text can be the same or different, dividing the encryption and decryption algorithms into the following types:

- **Symmetric key algorithm**—The keys for encryption and decryption are the same.
- **Asymmetric key algorithm**—The keys for encryption and decryption are different, one is the public key, and the other is the private key. The information encrypted with the public key can only be decrypted with the corresponding private key, and vice versa. The private key is kept secret, and the public key may be distributed widely. The private key cannot be practically derived from the public key. Asymmetric key algorithms include the Revest-Shamir-Adleman Algorithm (RSA), and the Digital Signature Algorithm (DSA).

Asymmetric key algorithms can be used in two scenarios for two purposes:

- **To encrypt and decrypt data**—The sender uses the public key of the intended receiver to encrypt the information to be sent. Only the intended receiver, the holder of the paired private key, can decrypt the information. This mechanism guarantees confidentiality. Only RSA can be used for data encryption and decryption.
- **To authenticate a sender**—Also called digital signature. The sender "signs" the information to be sent by encrypting the information with its own private key. A receiver decrypts the information with the sender's public key and, based on whether the information can be decrypted, determines the authenticity of the information. RSA and DSA can be used for digital signature.

Asymmetric key algorithms are widely used in various applications. For example, Secure Shell (SSH), Secure Sockets Layer (SSL), and Public Key Infrastructure (PKI) use the algorithms for digital signature. For information about SSH, SSL, and PKI, see "[Configuring SSH2.0](#)," "[Configuring SSL](#)," and "[Configuring PKI](#)."

## Configuration task list

Public key configuration tasks enable you to manage the local asymmetric key pairs, and configure the peer host public keys on the local device. By completing these tasks, the local device is ready to work with applications such as SSH and SSL to implement data encryption/decryption, or digital signature.

Complete these tasks to configure public keys:

| Task                                                        | Remarks                                                                                                                                                                             |
|-------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Configuring a local asymmetric key pair on the local device | <a href="#">Creating a local asymmetric key pair</a><br><a href="#">Displaying or exporting the local host public key</a><br><a href="#">Destroying a local asymmetric key pair</a> |
| Specifying the peer public key on the local device          | Optional                                                                                                                                                                            |

## Creating a local asymmetric key pair

When you create an asymmetric key pair on the local device, follow these guidelines:

- Create an asymmetric key pair of the proper type to work with a target application.
- After you enter the command, specify a proper modulus length for the key pair. The following table compares the three types of key pairs.

**Table 13 A comparison between different types of asymmetric key pairs**

| Type | Number of key pairs                                                                                              | Modulus length                      | Remarks                                              |
|------|------------------------------------------------------------------------------------------------------------------|-------------------------------------|------------------------------------------------------|
| RSA  | Two key pairs, one server key pair and one host key pair. Each key pair comprises a public key and a private key | 512 to 2048 bits<br>1024 by default | To achieve high security, specify at least 768 bits. |
| DSA  | One key pair, the host key pair                                                                                  |                                     |                                                      |

### ⚠ IMPORTANT:

Only SSH1.5 uses the RSA server key pair.

To create a local asymmetric key pair:

| Step                                   | Command                                      | Remarks                                                                                                                                                                     |
|----------------------------------------|----------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                  | <b>system-view</b>                           | N/A                                                                                                                                                                         |
| 2. Create a local asymmetric key pair. | <b>public-key local create { dsa   rsa }</b> | By default, no asymmetric key pair is created.<br>Key pairs created with the <b>public-key local create</b> command are saved automatically and can survive system reboots. |

## Displaying or exporting the local host public key

In SSH, to allow your local device to be authenticated by a peer device through digital signature, you must display or export the local host public key, which will then be specified on the peer device.

To display or export the local host public key, choose one of the following methods:

- [Displaying and recording the host public key information](#)
- [Displaying the host public key in a specific format and saving it to a file](#)
- [Exporting the host public key in a specific format to a file](#)

If your local device functions to authenticate the peer device, you must specify the peer public key on the local device. For more information, see "[Specifying the peer public key on the local device.](#)"

## Displaying and recording the host public key information

To display the local public key:

| Task                               | Command                                                                                                                             | Remarks                   |
|------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------|---------------------------|
| Display the local RSA public keys. | <b>display public-key local rsa public</b><br>[   { <b>begin</b>   <b>exclude</b>   <b>include</b> }<br><i>regular-expression</i> ] | Available in any view.    |
| Display the local host public key. | <b>display public-key local dsa public</b><br>[   { <b>begin</b>   <b>exclude</b>   <b>include</b> }<br><i>regular-expression</i> ] | Use at least one command. |

The **display public-key local rsa public** command displays both the RSA server and host public keys. Recording the RSA host public key is enough.

After displaying the host public key, record the key information for manual configuration of the key on the peer device.

## Displaying the host public key in a specific format and saving it to a file

To display the local host public key in a specific format:

| Step                                                                  | Command                                                                                                                                                                                                                                                                   | Remarks                   |
|-----------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------|
| 1. Enter system view.                                                 | <b>system-view</b>                                                                                                                                                                                                                                                        | N/A                       |
| 2. Display the local RSA or DSA host public key in a specific format. | <ul style="list-style-type: none"> <li>To display the local RSA host public key:<br/><b>public-key local export rsa { openssh   ssh1   ssh2 }</b></li> <li>To display the local DSA host public key:<br/><b>public-key local export dsa { openssh   ssh2 }</b></li> </ul> | Use at least one command. |

After you display the host public key in a specify format, save the key to a file, and transfer this file to the peer device.

## Exporting the host public key in a specific format to a file

After you export and save the host public key in a specify format to a file, transfer the file to the peer device.

To export and save the local host public key to a file:

| Step                                                                         | Command                                                                                                                                                                                                                                                                               | Remarks                   |
|------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------|
| 1. Enter system view.                                                        | <b>system-view</b>                                                                                                                                                                                                                                                                    | N/A                       |
| 2. Export a local RSA or DSA host public key in a specific format to a file. | <ul style="list-style-type: none"> <li>To export a local RSA host public key:<br/><b>public-key local export rsa { openssh   ssh1   ssh2 } filename</b></li> <li>To export a local DSA host public key:<br/><b>public-key local export dsa { openssh   ssh2 } filename</b></li> </ul> | Use at least one command. |



# Destroying a local asymmetric key pair

You may need to destroy a local asymmetric key pair and generate a new pair when an intrusion event has occurred, the storage media of the device is replaced, the asymmetric key has been used for a long time, or the local certificate expires. For more information about the local certificate, see "[Configuring PKI](#)."

To destroy a local asymmetric key pair:

| Step                                    | Command                                       |
|-----------------------------------------|-----------------------------------------------|
| 1. Enter system view.                   | <b>system-view</b>                            |
| 2. Destroy a local asymmetric key pair. | <b>public-key local destroy { dsa   rsa }</b> |

# Specifying the peer public key on the local device

In SSH, to enable the local device to authenticate a peer device, specify the peer public key on the local device. The device supports up to 20 peer public keys.

For information about displaying or exporting the host public key, see "[Displaying or exporting the local host public key](#)."

Take one of the following methods to specify the peer public key on the local device:

| Method                                                       | Prerequisites                                                                                                                                                                                                                                                                                                                                          | Remarks                                                                                                                                                                                                                                                                               |
|--------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Import the public key from a public key file (recommended)   | <ol style="list-style-type: none"><li>Save the host public key of the intended asymmetric key pair in a file.</li><li>Transfer a copy of the file through FTP or TFTP in binary mode to the local device.</li></ol>                                                                                                                                    | During the import process, the system automatically converts the public key to a string in Public Key Cryptography Standards (PKCS) format.                                                                                                                                           |
| Manually configure the public key—input or copy the key data | <ul style="list-style-type: none"><li>Display and record the public key of the intended asymmetric key pair.</li><li>If the peer device is an HP device, use the <b>display public-key local public</b> command to view and record its public key. A public key displayed by other methods for the HP device may not be in a correct format.</li></ul> | <ul style="list-style-type: none"><li>The recorded public key must be in the correct format, or the manual configuration of a format-incompliant public key will fail.</li><li>Always use the first method if you are not sure about the format of the recorded public key.</li></ul> |

To import the host public key from a public key file to the local device:

| Step                                                    | Command                                               |
|---------------------------------------------------------|-------------------------------------------------------|
| 1. Enter system view.                                   | <b>system-view</b>                                    |
| 2. Import the host public key from the public key file. | <b>public-key peer keyname import sshkey filename</b> |

To manually configure the peer public key on the local device:

| Step                                                            | Command                               | Remarks                                                                            |
|-----------------------------------------------------------------|---------------------------------------|------------------------------------------------------------------------------------|
| 1. Enter system view.                                           | <b>system-view</b>                    | N/A                                                                                |
| 2. Specify a name for the public key and enter public key view. | <b>public-key peer</b> <i>keyname</i> | N/A                                                                                |
| 3. Enter public key code view.                                  | <b>public-key-code begin</b>          | N/A                                                                                |
| 4. Configure the peer public key.                               | Type or copy the key                  | Spaces and carriage returns are allowed between characters.                        |
| 5. Return to public key view.                                   | <b>public-key-code end</b>            | When you exit public key code view, the system automatically saves the public key. |
| 6. Return to system view.                                       | <b>peer-public-key end</b>            | N/A                                                                                |

## Displaying and maintaining public keys

| Task                                                               | Command                                                                                                                                                                | Remarks               |
|--------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Display the local public keys.                                     | <b>display public-key local</b> { <i>dsa</i>   <i>rsa</i> } <b>public</b> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]           | Available in any view |
| Display the specified or all peer public keys on the local device. | <b>display public-key peer</b> [ <b>brief</b>   <b>name</b> <i>publickey-name</i> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] | Available in any view |

## Public key configuration examples

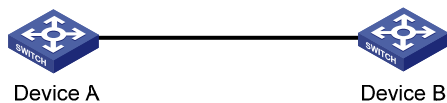
### Manually specifying the peer public key on the local device

#### Network requirements

As shown in [Figure 92](#), to prevent illegal access, Device B (the local device) authenticates Device A (the peer device) through a digital signature. Before configuring authentication parameters on Device B, configure the public key of Device A on Device B.

- Configure Device B to use the asymmetric key algorithm of RSA to authenticate Device A.
- Manually specify the host public key of Device A's public key pair on Device B.

**Figure 92 Network diagram**



#### Configuration procedure

1. Configure Device A;  

```
Create local RSA key pairs on Device A, setting the modulus length to the default, 1024 bits.
<DeviceA> system-view
```

```
[DeviceA] public-key local create rsa
The range of public key size is (512 ~ 2048).
NOTES: If the key modulus is greater than 512,
It will take a few minutes.
```

```
Press CTRL+C to abort.
```

```
Input the bits of the modulus[default = 1024]:
```

```
Generating Keys...
```

```
++++++
++++++
+++++++
+++++++
```

```
Display the public keys of the local RSA key pairs.
```

```
[DeviceA] display public-key local rsa public
```

```
=====
```

```
Time of Key pair created: 09:50:06 2012/03/07
```

```
Key name: HOST_KEY
```

```
Key type: RSA Encryption Key
```

```
=====
```

```
Key code:
```

```
30819F300D06092A864886F70D010101050003818D0030818902818100D90003FA95F5A44A2A2CD3F
814F9854C4421B57CAC64CFFE4782A87B0360B600497D87162D1F398E6E5E51E5E353B3A9AB16C9E7
66BD995C669A784AD597D0FB3AA9F7202C507072B19C3C50A0D7AD3994E14ABC62DB125035EA32647
0034DC078B2BAA3BC3BCA80AAB5EE01986BD1EF64B42F17CCAE4A77F1EF999B2BF9C4A10203010001
```

```
=====
```

```
Time of Key pair created: 09:50:07 2012/03/07
```

```
Key name: SERVER_KEY
```

```
Key type: RSA Encryption Key
```

```
=====
```

```
Key code:
```

```
307C300D06092A864886F70D0101010500036B003068026100999089E7AEE9802002D9EB2D0433B87
BB6158E35000AFB3FF310E42F109829D65BF70F7712507BE1A3E0BC5C2C03FAAF00DFDDC63D004B44
90DACBA3CFA9E84B9151BDC7EECE1C8770D961557D192DE2B36CAF9974B7B293363BB372771C2C1F0
203010001
```

## 2. Configure Device B:

# Configure the host public key of Device A's RSA key pairs on Device B. In public key code view, input the host public key of Device A. The host public key is the content of HOST\_KEY displayed on Device A by using the **display public-key local dsa public** command.

```
<DeviceB> system-view
```

```
[DeviceB] public-key peer devicea
```

```
Public key view: return to System View with "peer-public-key end".
```

```
[DeviceB-pkey-public-key] public-key-code begin
```

```
Public key code view: return to last view with "public-key-code end".
```

```
[DeviceB-pkey-key-code]30819F300D06092A864886F70D010101050003818D0030818902818100
D90003FA95F5A44A2A2CD3F814F9854C4421B57CAC64CFFE4782A87B0360B600497D87162D1F398E6
E5E51E5E353B3A9AB16C9E766BD995C669A784AD597D0FB3AA9F7202C507072B19C3C50A0D7AD3994
E14ABC62DB125035EA326470034DC078B2BAA3BC3BCA80AAB5EE01986BD1EF64B42F17CCAE4A77F1E
F999B2BF9C4A10203010001
```

```
[DeviceB-pkey-key-code] public-key-code end
[DeviceB-pkey-public-key] peer-public-key end
Display the host public key of Device A saved on Device B.
[DeviceB] display public-key peer name devicea
```

```
=====
Key Name : devicea
Key Type : RSA
Key Module: 1024
=====
Key Code:
30819F300D06092A864886F70D010101050003818D0030818902818100D90003FA95F5A44A2A2CD3F
814F9854C4421B57CAC64CFE4782A87B0360B600497D87162D1F398E6E5E51E5E353B3A9AB16C9E7
66BD995C669A784AD597D0FB3AA9F7202C507072B19C3C50A0D7AD3994E14ABC62DB125035EA32647
0034DC078B2BAA3BC3BCA80AAB5EE01986BD1EF64B42F17CCAE4A77F1EF999B2BF9C4A10203010001
```

The output shows that the host public key of Device A saved on Device B is consistent with the one created on Device A.

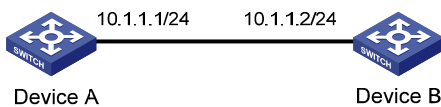
## Importing a peer public key from a public key file

### Network requirements

As shown in [Figure 93](#), to prevent illegal access, Device B (the local device) authenticates Device A (the peer device) through a digital signature. Before configuring authentication parameters on Device B, configure the public key of Device A on Device B.

- Configure Device B to use the asymmetric key algorithm of RSA to authenticate Device A.
- Import the host public key of Device A from the public key file to Device B.

**Figure 93 Network diagram**



### Configuration procedure

1. Create key pairs on Device A and export the host public key:

# Create local RSA key pairs on Device A, setting the modulus length to the default, 1024 bits.

```
<DeviceA> system-view
[DeviceA] public-key local create rsa
The range of public key size is (512 ~ 2048).
NOTES: If the key modulus is greater than 512,
It will take a few minutes.
Press CTRL+C to abort.
Input the bits of the modulus[default = 1024]:
Generating Keys...
++++++
++++++
+++++++
+++++++
```

```
Display the public keys of the local RSA key pairs.
```

```
[DeviceA] display public-key local rsa public
```

```
=====
```

```
Time of Key pair created: 09:50:06 2012/03/07
```

```
Key name: HOST_KEY
```

```
Key type: RSA Encryption Key
```

```
=====
```

```
Key code:
```

```
30819F300D06092A864886F70D010101050003818D0030818902818100D90003FA95F5A44A2A2CD3F
814F9854C4421B57CAC64CFE4782A87B0360B600497D87162D1F398E6E5E51E5E353B3A9AB16C9E7
66BD995C669A784AD597D0FB3AA9F7202C507072B19C3C50A0D7AD3994E14ABC62DB125035EA32647
0034DC078B2BAA3BC3BCA80AAB5EE01986BD1EF64B42F17CCA4E4A77F1EF999B2BF9C4A10203010001
```

```
=====
```

```
Time of Key pair created: 09:50:07 2012/03/07
```

```
Key name: SERVER_KEY
```

```
Key type: RSA Encryption Key
```

```
=====
```

```
Key code:
```

```
307C300D06092A864886F70D0101010500036E003068026100999089E7AEE9802002D9EB2D0433B87
BB6158E35000AFB3FF310E42F109829D65BF70F7712507BE1A3E0BC5C2C03FAAF00DFDDC63D004B44
90DACBA3CFA9E84B9151BDC7EECE1C8770D961557D192DE2B36CAF9974B7B293363BB372771C2C1F0
203010001
```

```
Export the RSA host public key HOST_KEY to a file named devicea.pub.
```

```
[DeviceA] public-key local export rsa ssh2 devicea.pub
```

2. Enable the FTP server function on Device A:

```
Enable the FTP server function, create an FTP user with the username ftp, password 123, and user level 3. This user level guarantees that the user has the permission to perform FTP operations.
```

```
[DeviceA] ftp server enable
```

```
[DeviceA] local-user ftp
```

```
[DeviceA-luser-ftp] password simple 123
```

```
[DeviceA-luser-ftp] service-type ftp
```

```
[DeviceA-luser-ftp] authorization-attribute level 3
```

```
[DeviceA-luser-ftp] quit
```

3. On Device B, get the public key file of Device A:

```
From Device B, use FTP to log in to Device A, and get the public key file devicea.pub with the file transfer mode of binary.
```

```
<DeviceB> ftp 10.1.1.1
```

```
Trying 10.1.1.1 ...
```

```
Press CTRL+K to abort
```

```
Connected to 10.1.1.1.
```

```
220 FTP service ready.
```

```
User(10.1.1.1:(none)):ftp
```

```
331 Password required for ftp.
```

```
Password:
```

```
230 User logged in.
```

```
[ftp] binary
```

```

200 Type set to I.
[ftp] get devicea.pub
227 Entering Passive Mode (10,1,1,1,5,148).
125 BINARY mode data connection already open, transfer starting for /devicea.pub.
226 Transfer complete.
FTP: 299 byte(s) received in 0.189 second(s), 1.00Kbyte(s)/sec.
[ftp] quit
221 Server closing.

```

4. Import the host public key of Device A to Device B:

# Import the host public key of Device A from the key file **devicea.pub** to Device B.

```

<DeviceB> system-view
[DeviceB] public-key peer devicea import sshkey devicea.pub

```

# Display the host public key of Device A on Device B.

```

[DeviceB] display public-key peer name devicea

```

```

=====
Key Name : devicea
Key Type : RSA
Key Module: 1024
=====
Key Code:
30819F300D06092A864886F70D010101050003818D0030818902818100D90003FA95F5A44A2A2CD3F
814F9854C4421B57CAC64CFE4782A87B0360B600497D87162D1F398E6E5E51E5E353B3A9AB16C9E7
66BD995C669A784AD597D0FB3AA9F7202C507072B19C3C50A0D7AD3994E14ABC62DB125035EA32647
0034DC078B2BAA3BC3BCA80AAB5EE01986BD1EF64B42F17CCAE4A77F1EF999B2BF9C4A10203010001

```

The output shows that the host public key of Device A saved on Device B is consistent with the one created on Device A.

---

# Configuring PKI

## Overview

The Public Key Infrastructure (PKI) is a general security infrastructure used to provide information security through public key technologies.

PKI, also called asymmetric key infrastructure, uses a key pair to encrypt and decrypt the data. The key pair consists of a private key and a public key. The private key must be kept secret but the public key needs to be distributed. Data encrypted by one of the two keys can only be decrypted by the other.

A key problem with PKI is how to manage the public keys. PKI employs the digital certificate mechanism to solve this problem. The digital certificate mechanism binds public keys to their owners, helping distribute public keys in large networks securely.

With digital certificates, the PKI system provides network communication and e-commerce with security services such as user authentication, data non-repudiation, data confidentiality, and data integrity.

HP's PKI system provides certificate management for Secure Sockets Layer (SSL).

## PKI terms

- Digital certificate

A digital certificate is a file signed by a certificate authority (CA) for an entity. It includes mainly the identity information of the entity, the public key of the entity, the name and signature of the CA, and the validity period of the certificate. The signature of the CA ensures the validity and authority of the certificate. A digital certificate must comply with the international standard of ITU-T X.509. The most common standard is X.509 v3.

This document discusses two types of certificates: local certificate and CA certificate. A local certificate is a digital certificate signed by a CA for an entity. A CA certificate is the certificate of a CA. If multiple CAs are trusted by different users in a PKI system, the CAs will form a CA tree with the root CA at the top level. The root CA has a CA certificate signed by itself and each lower level CA has a CA certificate signed by the CA at the next higher level.

- CRL

An existing certificate might need to be revoked when, for example, the username changes, the private key leaks, or the user stops the business. Revoking a certificate removes the binding of the public key with the user identity information. In PKI, the revocation is made through certificate revocation lists (CRLs). Whenever a certificate is revoked, the CA publishes one or more CRLs to show all certificates that have been revoked. The CRLs contain the serial numbers of all revoked certificates and provide an effective way for checking the validity of certificates.

A CA might publish multiple CRLs when the number of revoked certificates is so large that publishing them in a single CRL might degrade network performance. A CA uses CRL distribution points to indicate the URLs of these CRLs.

- CA policy

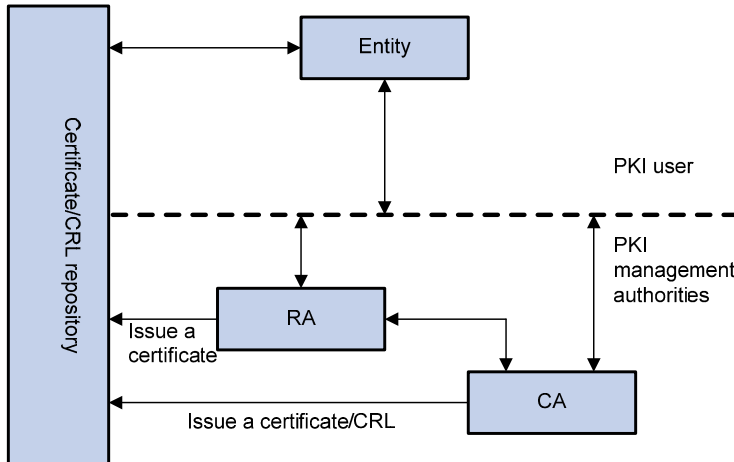
A CA policy is a set of criteria that a CA follows in processing certificate requests, issuing and revoking certificates, and publishing CRLs. Usually, a CA advertises its policy in the form of certification practice statement (CPS). A CA policy can be acquired through out-of-band means

such as phone, disk, and email. As different CAs might use different methods to examine the binding of a public key with an entity, make sure that you understand the CA policy before selecting a trusted CA for certificate request.

## PKI architecture

A PKI system consists of entities, a CA, a registration authority (RA) and a PKI repository.

**Figure 94 PKI architecture**



- Entity
 

An entity is an end user of PKI products or services, such as a person, an organization, a device, or a process running on a computer.
- CA
 

A CA is a trusted authority responsible for issuing and managing digital certificates. A CA issues certificates, specifies the validity periods of certificates, and revokes certificates as needed by publishing CRLs.
- RA
 

A registration authority (RA) is an extended part of a CA or an independent authority. An RA can implement functions including identity authentication, CRL management, key pair generation and key pair backup. The PKI standard recommends that an independent RA be used for registration management to achieve higher security.
- PKI repository
 

A PKI repository can be a Lightweight Directory Access Protocol (LDAP) server or a common database. It stores and manages information like certificate requests, certificates, keys, CRLs and logs when it provides a simple query function.

LDAP is a protocol for accessing and managing PKI information. An LDAP server stores user information and digital certificates from the RA server and provides directory navigation service. From an LDAP server, an entity can retrieve local and CA certificates of its own as well as certificates of other entities.

## PKI applications

The PKI technology can satisfy the security requirements of online transactions. As an infrastructure, PKI has a wide range of applications. Here are some application examples.



- **VPN**  
A virtual private network (VPN) is a private data communication network built on the public communication infrastructure. A VPN can leverage network layer security protocols (for instance, IPsec) in conjunction with PKI-based encryption and digital signature technologies for confidentiality.
- **Secure email**  
Emails require confidentiality, integrity, authentication, and non-repudiation. PKI can address these needs. The secure email protocol that is developing rapidly is Secure/Multipurpose Internet Mail Extensions (S/MIME), which is based on PKI and allows for transfer of encrypted mails with signature.
- **Web security**  
For web security, two peers can establish an SSL connection first for transparent and secure communications at the application layer. With PKI, SSL enables encrypted communications between a browser and a server. Both of the communication parties can verify each other's identity through digital certificates.

## How PKI operates

In a PKI-enabled network, an entity can request a local certificate from the CA and the device can check the validity of certificates. Here is how it operates:

1. An entity submits a certificate request to the RA.
2. The RA reviews the identity of the entity and then sends the identity information and the public key with a digital signature to the CA.
3. The CA verifies the digital signature, approves the application, and issues a certificate.
4. The RA receives the certificate from the CA, sends it to the LDAP server to provide directory navigation service, and notifies the entity that the certificate is successfully issued.
5. The entity retrieves the certificate. With the certificate, the entity can communicate with other entities safely through encryption and digital signature.
6. The entity makes a request to the CA when it needs to revoke its certificate. The CA approves the request, updates the CRLs and publishes the CRLs on the LDAP server.

## PKI configuration task list

| Task                                                                                                                                                                                                                                              | Remarks                           |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------|
| <a href="#">Configuring an entity DN</a>                                                                                                                                                                                                          | Required.                         |
| <a href="#">Configuring a PKI domain</a>                                                                                                                                                                                                          | Required.                         |
| <a href="#">Submitting a PKI certificate request</a> <ul style="list-style-type: none"> <li>• <a href="#">Submitting a certificate request in auto mode</a></li> <li>• <a href="#">Submitting a certificate request in manual mode</a></li> </ul> | Required.<br>Use either approach. |
| <a href="#">Retrieving a certificate manually</a>                                                                                                                                                                                                 | Optional.                         |
| <a href="#">Configuring PKI certificate verification</a>                                                                                                                                                                                          | Optional.                         |
| <a href="#">Destroying a local RSA key pair</a>                                                                                                                                                                                                   | Optional.                         |
| <a href="#">Deleting a certificate</a>                                                                                                                                                                                                            | Optional.                         |

| Task                                                 | Remarks   |
|------------------------------------------------------|-----------|
| <a href="#">Configuring an access control policy</a> | Optional. |

## Configuring an entity DN

A certificate is the binding of a public key and the identity information of an entity, where the identity information is identified by an entity distinguished name (DN). A CA identifies a certificate applicant uniquely by entity DN.

An entity DN is defined by these parameters:

- Common name of the entity.
- Country code of the entity, a standard 2-character code. For example, CN represents China and US represents the United States.
- Fully qualified domain name (FQDN) of the entity, a unique identifier of an entity on the network. It consists of a host name and a domain name and can be resolved to an IP address. For example, **www.whatever.com** is an FQDN, where **www** is a host name and **whatever.com** a domain name.
- IP address of the entity.
- Locality where the entity resides.
- Organization to which the entity belongs.
- Unit of the entity in the organization.
- State where the entity resides.

The configuration of an entity DN must comply with the CA certificate issue policy. You must determine, for example, which entity DN parameters are mandatory and which are optional. Otherwise, certificate requests might be rejected.

To configure an entity DN:

| Step                                          | Command                                | Remarks                                               |
|-----------------------------------------------|----------------------------------------|-------------------------------------------------------|
| 1. Enter system view.                         | <b>system-view</b>                     | N/A                                                   |
| 2. Create an entity and enter its view.       | <b>pki entity</b> <i>entity-name</i>   | No entity exists by default.                          |
| 3. Configure the common name for the entity.  | <b>common-name</b> <i>name</i>         | Optional.<br>No common name is specified by default.  |
| 4. Configure the country code for the entity. | <b>country</b> <i>country-code-str</i> | Optional.<br>No country code is specified by default. |
| 5. Configure the FQDN for the entity.         | <b>fqdn</b> <i>name-str</i>            | Optional.<br>No FQDN is specified by default.         |
| 6. Configure the IP address for the entity.   | <b>ip</b> <i>ip-address</i>            | Optional.<br>No IP address is specified by default.   |
| 7. Configure the locality for the entity.     | <b>locality</b> <i>locality-name</i>   | Optional.<br>No locality is specified by default.     |

| Step                                                | Command                                       | Remarks                                                    |
|-----------------------------------------------------|-----------------------------------------------|------------------------------------------------------------|
| 8. Configure the organization name for the entity.  | <b>organization</b> <i>org-name</i>           | Optional.<br>No organization is specified by default.      |
| 9. Configure the unit name for the entity.          | <b>organization-unit</b> <i>org-unit-name</i> | Optional.<br>No unit is specified by default.              |
| 10. Configure the state or province for the entity. | <b>state</b> <i>state-name</i>                | Optional.<br>No state or province is specified by default. |

#### NOTE:

The Windows 2000 CA server has some restrictions on the data length of a certificate request. If the entity DN in a certificate request goes beyond a certain limit, the server will not respond to the certificate request.

## Configuring a PKI domain

Before requesting a PKI certificate, an entity needs to be configured with some enrollment information, which is referred to as a PKI domain. A PKI domain is only intended for convenient reference by applications like SSL, and only has local significance. A PKI domain configured on a switch is invisible to the CA and other switches, and each PKI domain has its own parameters.

A PKI domain defines these parameters:

- **Trusted CA**—An entity requests a certificate from a trusted CA.
- **Entity**—A certificate applicant uses an entity to provide its identity information to a CA.
- **RA**—Generally, an independent RA is in charge of certificate request management. It receives the registration request from an entity, examines its qualification, and determines whether to ask the CA to sign a digital certificate. The RA only examines the application qualification of an entity; it does not issue any certificate. Sometimes, the registration management function is provided by the CA, in which case no independent RA is required. It is a good practice to deploy an independent RA.
- **URL of the registration server**—An entity sends a certificate request to the registration server through Simple Certification Enrollment Protocol (SCEP), a dedicated protocol for an entity to communicate with a CA. This URL is also called the certificate request URL.
- **Polling interval and count**—After an applicant makes a certificate request, the CA might need a long period of time if it verifies the certificate request manually. During this period, the applicant needs to query the status of the request periodically to get the certificate as soon as possible after the certificate is signed. You can configure the polling interval and count to query the request status.
- **IP address of the LDAP server**—An LDAP server is usually deployed to store certificates and CRLs. If this is the case, you must configure the IP address of the LDAP server.
- **Fingerprint for root certificate verification**—After receiving the root certificate of the CA, an entity needs to verify the fingerprint of the root certificate, namely, the hash value of the root certificate content. This hash value is unique to every certificate. If the fingerprint of the root certificate does not match the one configured for the PKI domain, the entity will reject the root certificate.

## Configuration guidelines

- Up to two PKI domains can be created on a switch.

- The CA name is required only when you retrieve a CA certificate. It is not used when in local certificate request.
- The certificate request URL does not support domain name resolution.

## Configuration procedure

To configure a PKI domain:

| Step                                                                                             | Command                                                                                                             | Remarks                                                                                                                                                                                                                                                                               |
|--------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                                                                            | <b>system-view</b>                                                                                                  | N/A                                                                                                                                                                                                                                                                                   |
| 2. Create a PKI domain and enter its view.                                                       | <b>pki domain</b> <i>domain-name</i>                                                                                | No PKI domain exists by default.                                                                                                                                                                                                                                                      |
| 3. Specify the trusted CA.                                                                       | <b>ca identifier</b> <i>name</i>                                                                                    | No trusted CA is specified by default.                                                                                                                                                                                                                                                |
| 4. Specify the entity for certificate request.                                                   | <b>certificate request entity</b><br><i>entity-name</i>                                                             | No entity is specified by default.<br>The specified entity must exist.                                                                                                                                                                                                                |
| 5. Specify the authority for certificate request.                                                | <b>certificate request from</b> { <b>ca</b>   <b>ra</b> }                                                           | No authority is specified by default.                                                                                                                                                                                                                                                 |
| 6. Configure the certificate request URL.                                                        | <b>certificate request url</b> <i>url-string</i>                                                                    | No certificate request URL is configured by default.                                                                                                                                                                                                                                  |
| 7. Configure the polling interval and attempt limit for querying the certificate request status. | <b>certificate request polling</b> { <b>count</b> <i>count</i>   <b>interval</b> <i>minutes</i> }                   | Optional.<br>The polling is executed for up to 50 times at the interval of 20 minutes by default.                                                                                                                                                                                     |
| 8. Specify the LDAP server.                                                                      | <b>ldap-server ip</b> <i>ip-address</i> [ <b>port</b> <i>port-number</i> ] [ <b>version</b> <i>version-number</i> ] | Optional.<br>No LDP server is specified by default.                                                                                                                                                                                                                                   |
| 9. Configure the fingerprint for root certificate verification.                                  | <b>root-certificate fingerprint</b> { <b>md5</b>   <b>sha1</b> } <i>string</i>                                      | Required when the certificate request mode is auto and optional when the certificate request mode is manual. In the latter case, if you do not configure this command, the fingerprint of the root certificate must be verified manually.<br>No fingerprint is configured by default. |

## Submitting a PKI certificate request

When requesting a certificate, an entity introduces itself to the CA by providing its identity information and public key, which will be the major components of the certificate. A certificate request can be submitted to a CA in offline mode or online mode. In offline mode, a certificate request is submitted to a CA by an "out-of-band" means such as phone, disk, or email.

An online certificate request can be submitted in manual mode or auto mode.

## Submitting a certificate request in auto mode

In auto mode, an entity automatically requests a certificate from the CA server if it has no local certificate for an application working with PKI, and then retrieves the certificate and saves the certificate locally. Before requesting a certificate, if the PKI domain does not have the CA certificate yet, the entity automatically retrieves the CA certificate.

To configure an entity to submit a certificate request in auto mode:

| Step                                         | Command                                                                                                                                               | Remarks           |
|----------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|
| 1. Enter system view.                        | <b>system-view</b>                                                                                                                                    | N/A               |
| 2. Enter PKI domain view.                    | <b>pki domain</b> <i>domain-name</i>                                                                                                                  | N/A               |
| 3. Set the certificate request mode to auto. | <b>certificate request mode auto</b><br>[ <b>key-length</b> <i>key-length</i>   <b>password</b> { <b>cipher</b>   <b>simple</b> } <i>password</i> ] * | Manual by default |

### ! IMPORTANT:

In auto mode, an entity does not automatically re-request a certificate to replace a certificate that is expiring or has expired. After the certificate expires, the service using the certificate might be interrupted.

## Submitting a certificate request in manual mode

In manual mode, you manually submit a certificate request for an entity. Before submitting a certificate request, you must make sure that an RSA key pair has been generated and the CA certificate has been retrieved and saved locally.

The CA certificate is required to verify the authenticity and validity of a local certificate. The public key of the key pair is an important part of the request information and will be transferred to the CA along with some other information. For more information about RSA key pair configuration, see *Security Configuration Guide*.

### Configuration guidelines

- If a PKI domain already has a local certificate, creating an RSA key pair will result in inconsistency between the key pair and the certificate. To generate a new RSA key pair, delete the local certificate and then issue the **public-key local create** command. For more information about the **public-key local create** command, see *Security Command Reference*.
- A newly created key pair will overwrite the existing one. If you perform the **public-key local create** command in the presence of a local RSA key pair, the system will ask you whether you want to overwrite the existing one.
- If a PKI domain already has a local certificate, you cannot request another certificate for it. This helps avoid inconsistency between the certificate and the registration information resulting from configuration changes. Before requesting a new certificate, use the **pki delete-certificate** command to delete the existing local certificate and the CA certificate stored locally.
- When it is impossible to request a certificate from the CA through SCEP, you can print the request information or save the request information to a local file, and then send the printed information or saved file to the CA by an out-of-band means. To print the request information, use the **pki request-certificate domain** command with the **pkcs10** keyword. To save the request information to a local file, use the **pki request-certificate domain** command with the **pkcs10 filename** *filename* option.

- Make sure the clocks of the entity and the CA are synchronous. Otherwise, the validity period of the certificate will be abnormal.
- The configuration made by the **pki request-certificate domain** command is not saved in the configuration file.

## Configuration procedure

To submit a certificate request in manual mode:

| Step                                            | Command                                                                                                                           | Remarks                                  |
|-------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------|------------------------------------------|
| 1. Enter system view.                           | <b>system-view</b>                                                                                                                | N/A                                      |
| 2. Enter PKI domain view.                       | <b>pki domain</b> <i>domain-name</i>                                                                                              | N/A                                      |
| 3. Set the certificate request mode to manual.  | <b>certificate request mode manual</b>                                                                                            | Optional.<br>Manual by default.          |
| 4. Return to system view.                       | <b>quit</b>                                                                                                                       | N/A                                      |
| 5. Retrieve a CA certificate manually.          | See " <a href="#">Retrieving a certificate manually</a> "                                                                         | N/A                                      |
| 6. Generate a local RSA key pair.               | <b>public-key local create rsa</b>                                                                                                | No local RSA key pair exists by default. |
| 7. Submit a local certificate request manually. | <b>pki request-certificate domain</b><br><i>domain-name</i> [ <i>password</i> ]<br>[ <b>pkcs10</b> [ <i>filename filename</i> ] ] | N/A                                      |

## Retrieving a certificate manually

You can download CA certificates, local certificates, or peer entity certificates from the CA server and save them locally. To do so, use either the offline mode or the online mode. In offline mode, you must retrieve a certificate by an out-of-band means like FTP, disk, or email, and then import it into the local PKI system.

Certificate retrieval serves the following purposes:

- Locally store the certificates associated with the local security domain for improved query efficiency and reduced query count
- Prepare for certificate verification

## Configuration guidelines

- Before retrieving a local certificate in online mode, be sure to complete the LDAP server configuration.
- If a PKI domain already has a CA certificate, you cannot retrieve another CA certificate for it. This restriction helps avoid inconsistency between the certificate and registration information resulted from configuration changes. To retrieve a new CA certificate, use the **pki delete-certificate** command to delete the existing CA certificate and the local certificate first.
- The configuration made by the **pki retrieval-certificate** configuration is not saved in the configuration file.
- Make sure the switch's system time falls in the validity period of the certificate so that the certificate is valid.

## Configuration procedure

To retrieve a certificate manually:

| Step                                | Command                                                                                                                                                                                                                                                                                                                                                                              | Remarks             |
|-------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| 1. Enter system view.               | <b>system-view</b>                                                                                                                                                                                                                                                                                                                                                                   | N/A                 |
| 2. Retrieve a certificate manually. | <ul style="list-style-type: none"><li>In online mode:<br/><b>pki retrieval-certificate</b> { <b>ca</b>   <b>local</b> } <b>domain</b><br/><i>domain-name</i></li><li>In offline mode:<br/><b>pki import-certificate</b> { <b>ca</b>   <b>local</b> } <b>domain</b><br/><i>domain-name</i> { <b>der</b>   <b>p12</b>   <b>pem</b> } [ <b>filename</b><br/><i>filename</i> ]</li></ul> | Use either command. |

## Configuring PKI certificate verification

A certificate needs to be verified before being used. Certificate verification can examine whether the certificate is signed by the CA and whether the certificate has expired or been revoked.

You can specify whether to perform CRL checking during certificate verification. If you enable CRL checking, CRLs will be used in verification of a certificate, and you must retrieve the CA certificate and CRLs to the local switch before the certificate verification. If you disable CRL checking, you only need to retrieve the CA certificate.

## Configuration guidelines

- The CRL update period defines the interval at which the entity downloads CRLs from the CRL server. The CRL update period setting manually configured on the switch is prior to that carried in the CRLs.
- The configuration made by the **pki retrieval-crl domain** command is not saved in the configuration file.
- The URL of the CRL distribution point does not support domain name resolution.

## Configuring CRL-checking-enabled PKI certificate verification

| Step                                              | Command                               | Remarks                                                                                          |
|---------------------------------------------------|---------------------------------------|--------------------------------------------------------------------------------------------------|
| 1. Enter system view.                             | <b>system-view</b>                    | N/A                                                                                              |
| 2. Enter PKI domain view.                         | <b>pki domain</b> <i>domain-name</i>  | N/A                                                                                              |
| 3. Specify the URL of the CRL distribution point. | <b>crl url</b> <i>url-string</i>      | Optional.<br>No CRL distribution point URL is specified by default.                              |
| 4. Set the CRL update period.                     | <b>crl update-period</b> <i>hours</i> | Optional.<br>By default, the CRL update period depends on the next update field in the CRL file. |
| 5. Enable CRL checking.                           | <b>crl check enable</b>               | Optional.<br>Enabled by default.                                                                 |

| Step                                     | Command                                                                            | Remarks |
|------------------------------------------|------------------------------------------------------------------------------------|---------|
| 6. Return to system view.                | <b>quit</b>                                                                        | N/A     |
| 7. Retrieve the CA certificate.          | See " <a href="#">Retrieving a certificate manually</a> "                          | N/A     |
| 8. Retrieve CRLs.                        | <b>pki retrieval-crl domain</b><br><i>domain-name</i>                              | N/A     |
| 9. Verify the validity of a certificate. | <b>pki validate-certificate { ca   local }</b><br><b>domain</b> <i>domain-name</i> | N/A     |

## Configuring CRL-checking-disabled PKI certificate verification

To configure CRL-checking-disabled PKI certificate verification:

| Step                                       | Command                                                                            | Remarks            |
|--------------------------------------------|------------------------------------------------------------------------------------|--------------------|
| 1. Enter system view.                      | <b>system-view</b>                                                                 | N/A                |
| 2. Enter PKI domain view.                  | <b>pki domain</b> <i>domain-name</i>                                               | N/A                |
| 3. Disable CRL checking.                   | <b>crl check disable</b>                                                           | Enabled by default |
| 4. Return to system view.                  | <b>quit</b>                                                                        | N/A                |
| 5. Retrieve the CA certificate.            | See " <a href="#">Retrieving a certificate manually</a> "                          | N/A                |
| 6. Verify the validity of the certificate. | <b>pki validate-certificate { ca   local }</b><br><b>domain</b> <i>domain-name</i> | N/A                |

## Destroying a local RSA key pair

A certificate has a lifetime, which is determined by the CA. When the private key leaks or the certificate is about to expire, you can destroy the old RSA key pair and then create a pair to request a new certificate.

To destroy a local RSA key pair:

| Step                             | Command                             |
|----------------------------------|-------------------------------------|
| 1. Enter system view.            | <b>system-view</b>                  |
| 2. Destroy a local RSA key pair. | <b>public-key local destroy rsa</b> |

For more information about the **public-key local destroy** command, see *Security Command Reference*.

## Deleting a certificate

When a certificate requested manually is about to expire or you want to request a new certificate, you can delete the current local certificate or CA certificate.



To delete a certificate:

| Step                    | Command                                                                                     |
|-------------------------|---------------------------------------------------------------------------------------------|
| 1. Enter system view.   | <b>system-view</b>                                                                          |
| 2. Delete certificates. | <b>pki delete-certificate</b> { <b>ca</b>   <b>local</b> } <b>domain</b> <i>domain-name</i> |

## Configuring an access control policy

By configuring a certificate attribute-based access control policy, you can further control access to the server, providing additional security for the server.

To configure a certificate attribute-based access control policy:

| Step                                                                                                                   | Command                                                                                                                                                                                                                                                     | Remarks                                                                                                                  |
|------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                                                                                                  | <b>system-view</b>                                                                                                                                                                                                                                          | N/A                                                                                                                      |
| 2. Create a certificate attribute group and enter its view.                                                            | <b>pki certificate attribute-group</b><br><i>group-name</i>                                                                                                                                                                                                 | No certificate attribute group exists by default.                                                                        |
| 3. Configure an attribute rule for the certificate issuer name, certificate subject name, or alternative subject name. | <b>attribute</b> <i>id</i> { <b>alt-subject-name</b> { <b>fqdn</b>   <b>ip</b> }   { <b>issuer-name</b>   <b>subject-name</b> } { <b>dn</b>   <b>fqdn</b>   <b>ip</b> } } { <b>ctn</b>   <b>equ</b>   <b>nctn</b>   <b>nequ</b> }<br><i>attribute-value</i> | Optional.<br>No restriction exists on the issuer name, certificate subject name and alternative subject name by default. |
| 4. Return to system view.                                                                                              | <b>quit</b>                                                                                                                                                                                                                                                 | N/A                                                                                                                      |
| 5. Create a certificate attribute-based access control policy and enter its view.                                      | <b>pki certificate access-control-policy</b><br><i>policy-name</i>                                                                                                                                                                                          | No access control policy exists by default.                                                                              |
| 6. Configure a certificate attribute-based access control rule.                                                        | <b>rule</b> [ <i>id</i> ] { <b>deny</b>   <b>permit</b> }<br><i>group-name</i>                                                                                                                                                                              | No access control rule exists by default.<br>A certificate attribute group must exist to be associated with a rule.      |

## Displaying and maintaining PKI

| Task                                                     | Command                                                                                                                                                                                                       | Remarks               |
|----------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Display the contents or request status of a certificate. | <b>display pki certificate</b> { { <b>ca</b>   <b>local</b> } <b>domain</b> <i>domain-name</i>   <b>request-status</b> [ [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] ] | Available in any view |
| Display CRLs.                                            | <b>display pki crl</b> <b>domain</b> <i>domain-name</i> [ [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] ]                                                                | Available in any view |

| Task                                                                           | Command                                                                                                                                                                     | Remarks               |
|--------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Display information about certificate attribute groups.                        | <b>display pki certificate attribute-group</b> { <i>group-name</i>   <b>all</b> } [ [ { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]        | Available in any view |
| Display information about certificate attribute-based access control policies. | <b>display pki certificate access-control-policy</b> { <i>policy-name</i>   <b>all</b> } [ [ { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] | Available in any view |

## PKI configuration examples

This section describes details about PKI configuration examples.

When the CA uses Windows Server, the SCEP add-on is required, and you must use the **certificate request from ra** command to specify that the entity request a certificate from an RA.

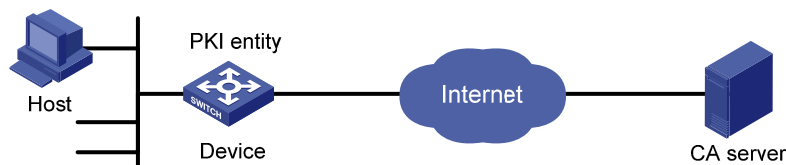
When the CA uses RSA Keon, the SCEP add-on is not required, and you must use the **certificate request from ca** command to specify that the entity request a certificate from a CA.

## Requesting a certificate from a CA server running RSA Keon

### Network requirements

The switch submits a local certificate request to the CA server. The switch acquires the CRLs for certificate verification.

Figure 95 Network diagram



### Configuring the CA server

1. Create a CA server named **myca**:

In this example, you need to configure these basic attributes on the CA server at first:

- **Nickname**—Name of the trusted CA.
- **Subject DN**—DN information of the CA, including the Common Name (CN), Organization Unit (OU), Organization (O), and Country (C).

Use the default values for the other attributes.

2. Configure extended attributes:

After configuring the basic attributes, perform configuration on the jurisdiction configuration page of the CA server. This includes selecting the proper extension profiles, enabling the SCEP autovetting function, and adding the IP address list for SCEP autovetting.

3. Configure the CRL distribution behavior:

After completing the configuration, you must perform CRL related configurations. In this example, select the local CRL distribution mode of Hypertext Transfer Protocol (HTTP) and set the HTTP URL to `http://4.4.4.133:447/myca.crl`.

After the configuration, make sure the system clock of the switch is synchronous to that of the CA, so that the switch can request certificates and retrieve CRLs properly.

## Configuring the switch

1. Configure the entity DN:

# Configure the entity name as **aaa** and the common name as **device**.

```
<Device> system-view
[Device] pki entity aaa
[Device-pki-entity-aaa] common-name device
[Device-pki-entity-aaa] quit
```

2. Configure the PKI domain:

# Create PKI domain **torsa** and enter its view.

```
[Device] pki domain torsa
```

# Configure the name of the trusted CA as **myca**.

```
[Device-pki-domain-torsa] ca identifier myca
```

# Configure the URL of the registration server in the format of `http://host:port/Issuing Jurisdiction ID`, where Issuing Jurisdiction ID is a hexadecimal string generated on the CA server.

```
[Device-pki-domain-torsa] certificate request url
http://4.4.4.133:446/c95e970f632d27be5e8cbf80e971d9c4a9a93337
```

# Set the registration authority to **CA**.

```
[Device-pki-domain-torsa] certificate request from ca
```

# Specify the entity for certificate request as **aaa**.

```
[Device-pki-domain-torsa] certificate request entity aaa
```

# Configure the URL for the CRL distribution point.

```
[Device-pki-domain-torsa] crl url http://4.4.4.133:447/myca.crl
[Device-pki-domain-torsa] quit
```

3. Generate a local key pair using RSA:

```
[Device] public-key local create rsa
```

The range of public key size is (512 ~ 2048).

NOTES: If the key modulus is greater than 512,

It will take a few minutes.

Press CTRL+C to abort.

Input the bits in the modulus [default = 1024]:

Generating Keys...

```
+++++
+++++
+++++
+++++
```

4. Apply for certificates:

# Retrieve the CA certificate and save it locally.

```
[Device] pki retrieval-certificate ca domain torsa
```

Retrieving CA/RA certificates. Please wait a while.....

```

The trusted CA's finger print is:
 MD5 fingerprint:EDE9 0394 A273 B61A F1B3 0072 A0B1 F9AB
 SHA1 fingerprint: 77F9 A077 2FB8 088C 550B A33C 2410 D354 23B2 73A8

Is the finger print correct?(Y/N):y

Saving CA/RA certificates chain, please wait a moment.....
CA certificates retrieval success.
Retrieve CRLs and save them locally.
[Device] pki retrieval-crl domain torsa
Connecting to server for retrieving CRL. Please wait a while.....
CRL retrieval success!
Request a local certificate manually.
[Device] pki request-certificate domain torsa challenge-word
Certificate is being requested, please wait.....
[Device]
Enrolling the local certificate,please wait a while.....
Certificate request Successfully!
Saving the local certificate to device.....
Done!

```

## Verifying the configuration

# Use the following command to view information about the local certificate acquired.

```

[Device] display pki certificate local domain torsa
Certificate:
Data:
 Version: 3 (0x2)
 Serial Number:
 9A96A48F 9A509FD7 05FFF4DF 104AD094
 Signature Algorithm: sha1WithRSAEncryption
 Issuer:
 C=cn
 O=org
 OU=test
 CN=myca
 Validity
 Not Before: Jan 8 09:26:53 2012 GMT
 Not After : Jan 8 09:26:53 2012 GMT
 Subject:
 CN=device
 Subject Public Key Info:
 Public Key Algorithm: rsaEncryption
 RSA Public Key: (1024 bit)
 Modulus (1024 bit):
 00D67D50 41046F6A 43610335 CA6C4B11
 F8F89138 E4E905BD 43953BA2 623A54C0
 EA3CB6E0 B04649CE C9CDDD38 34015970
 981E96D9 FF4F7B73 A5155649 E583AC61

```

```

D3A5C849 CBDE350D 2A1926B7 0AE5EF5E
D1D8B08A DBF16205 7C2A4011 05F11094
73EB0549 A65D9E74 0F2953F2 D4F0042F
19103439 3D4F9359 88FB59F3 8D4B2F6C
2B
Exponent: 65537 (0x10001)
X509v3 extensions:
 X509v3 CRL Distribution Points:
 URI:http://4.4.4.133:447/myca.crl

```

```

Signature Algorithm: sha1WithRSAEncryption
836213A4 F2F74C1A 50F4100D B764D6CE
B30C0133 C4363F2F 73454D51 E9F95962
EDE9E590 E7458FA6 765A0D3F C4047BC2
9C391FF0 7383C4DF 9A0CCFA9 231428AF
987B029C C857AD96 E4C92441 9382E798
8FCC1E4A 3E598D81 96476875 E2F86C33
75B51661 B6556C5E 8F546E97 5197734B
C8C29AC7 E427C8E4 B9AAF5AA 80A75B3C

```

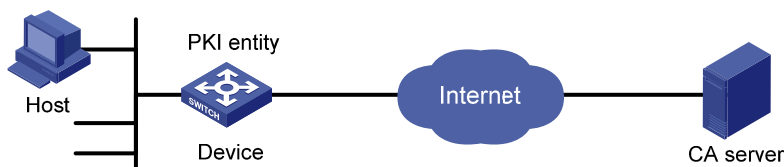
You can also use **display pki certificate ca domain** and **display pki crl domain** to display detailed information about the CA certificate and CRLs. For more information about the commands, see *Security Command Reference*.

## Requesting a certificate from a CA server running Windows 2003 Server

### Network requirements

Configure PKI entity Device to request a local certificate from the CA server.

Figure 96 Network diagram



### Configuring the CA server

1. Install the certificate service suites:
  - a. Select **Control Panel > Add or Remove Programs** from the start menu.
  - b. Select **Add/Remove Windows Components > Certificate Services**.
  - c. Click **Next** to begin the installation.
2. Install the SCEP add-on:

Because a CA server running the Windows 2003 server does not support SCEP by default, you must install the SCEP add-on so that the switch can register and obtain its certificate automatically. After the SCEP add-on installation completes, a URL is displayed, which you must configure on the switch as the URL of the server for certificate registration.

3. Modify the certificate service attributes:
  - a. Select **Control Panel > Administrative Tools > Certificate Authority** from the start menu.  
If the CA server and SCEP add-on have been installed successfully, there should be two certificates issued by the CA to the RA.
  - b. Right-click the CA server in the navigation tree and select **Properties > Policy Module**.
  - c. Click **Properties** and select **Follow the settings in the certificate template, if applicable. Otherwise, automatically issue the certificate**.
4. Modify the Internet Information Services (IIS) attributes:
  - a. Select **Control Panel > Administrative Tools > Internet Information Services (IIS) Manager** from the start menu.
  - b. Select **Web Sites** from the navigation tree.
  - c. Right-click **Default Web Site** and select **Properties > Home Directory**.
  - d. Specify the path for certificate service in the **Local path** text box.  
To avoid conflict with existing services, specify an available port number as the TCP port number of the default website.

After completing the configuration, make sure the system clock of the switch is synchronous to that of the CA server, so that that the switch can request a certificate normally.

## Configuring the switch

1. Configure the entity DN:
 

```
Configure the entity name as aaa and the common name as device.
<Device> system-view
[Device] pki entity aaa
[Device-pki-entity-aaa] common-name device
[Device-pki-entity-aaa] quit
```
2. Configure the PKI domain:
 

```
Create PKI domain torsa and enter its view.
[Device] pki domain torsa

Configure the name of the trusted CA as myca.
[Device-pki-domain-torsa] ca identifier myca

Configure the URL of the registration server in the format of http://host:port/
certsrv/mscep/mscep.dll, where host:port indicates the IP address and port number of the CA
server.
[Device-pki-domain-torsa] certificate request url
http://4.4.4.1:8080/certsrv/mscep/mscep.dll

Set the registration authority to RA.
[Device-pki-domain-torsa] certificate request from ra

Specify the entity for certificate request as aaa.
[Device-pki-domain-torsa] certificate request entity aaa
```
3. Generate a local key pair using RSA:
 

```
[Device] public-key local create rsa
The range of public key size is (512 ~ 2048).
NOTES: If the key modulus is greater than 512,
It will take a few minutes.
Press CTRL+C to abort.
```

```
Input the bits in the modulus [default = 1024]:
Generating Keys...
+++++
+++++
+++++
+++++
```

#### 4. Apply for certificates:

**# Retrieve the CA certificate and save it locally.**

```
[Device] pki retrieval-certificate ca domain torsa
Retrieving CA/RA certificates. Please wait a while.....
The trusted CA's finger print is:
 MD5 fingerprint:766C D2C8 9E46 845B 4DCE 439C 1C1F 83AB
 SHA1 fingerprint:97E5 DDED AB39 3141 75FB DB5C E7F8 D7D7 7C9B 97B4
```

```
Is the finger print correct?(Y/N):y
```

```
Saving CA/RA certificates chain, please wait a moment.....
CA certificates retrieval success.
```

**# Request a local certificate manually.**

```
[Device] pki request-certificate domain torsa challenge-word
Certificate is being requested, please wait.....
[Device]
Enrolling the local certificate,please wait a while.....
Certificate request Successfully!
Saving the local certificate to device.....
Done!
```

### Verifying the configuration

**# Use the following command to view information about the local certificate acquired.**

```
[Device] display pki certificate local domain torsa
Certificate:
 Data:
 Version: 3 (0x2)
 Serial Number:
 48FA0FD9 00000000 000C
 Signature Algorithm: sha1WithRSAEncryption
 Issuer:
 CN=myca
 Validity
 Not Before: Feb 21 12:32:16 2012 GMT
 Not After : Feb 21 12:42:16 2012 GMT
 Subject:
 CN=device
 Subject Public Key Info:
 Public Key Algorithm: rsaEncryption
 RSA Public Key: (1024 bit)
 Modulus (1024 bit):
```

```
00A6637A 8CDEA1AC B2E04A59 F7F6A9FE
5AEE52AE 14A392E4 E0E5D458 0D341113
0BF91E57 FA8C67AC 6CE8FE8B 5570178B
10242FDD D3947F5E 2DA70BD9 1FAF07E5
1D167CE1 FC20394F 476F5C08 C5067DF9
CB4D05E6 55DC11B6 9F4C014D EA600306
81D403CF 2D93BC5A 8AF3224D 1125E439
78ECEFE1 7FA9AE7B 877B50B8 3280509F
6B
```

```
Exponent: 65537 (0x10001)
```

```
X509v3 extensions:
```

```
X509v3 Subject Key Identifier:
```

```
B68E4107 91D7C44C 7ABCE3BA 9BF385F8 A448F4E1
```

```
X509v3 Authority Key Identifier:
```

```
keyid:9D823258 EADFEFA2 4A663E75 F416B6F6 D41EE4FE
```

```
X509v3 CRL Distribution Points:
```

```
URI:http://100192b/CertEnroll/CA%20server.crl
```

```
URI:file://\100192b\CertEnroll\CA server.crl
```

```
Authority Information Access:
```

```
CA Issuers - URI:http://100192b/CertEnroll/100192b_CA%20server.crt
```

```
CA Issuers - URI:file://\100192b\CertEnroll\100192b_CA server.crt
```

```
1.3.6.1.4.1.311.20.2:
```

```
.0.I.P.S.E.C.I.n.t.e.r.m.e.d.i.a.t.e.O.f.f.l.i.n.e
```

```
Signature Algorithm: sha1WithRSAEncryption
```

```
81029589 7BFA1CBD 20023136 B068840B
```

```
(Omitted)
```

You can also use some other **display** commands to display more information about the CA certificate. For more information about the **display pki certificate ca domain** command, see *Security Command Reference*.

## Configuring a certificate attribute-based access control policy

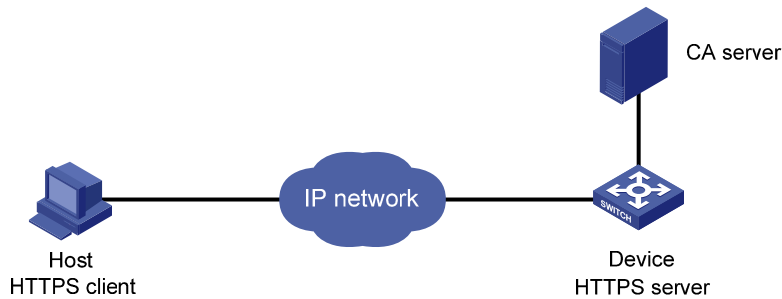
### Network requirements

The client accesses the remote HTTP Secure (HTTPS) server through the HTTPS protocol.

Configure SSL to make sure that only legal clients log into the HTTPS server, and create a certificate attribute-based access control policy to control access to the HTTPS server.



Figure 97 Network diagram



## Configuration procedure

The configuration procedure involves SSL configuration and HTTPS configuration. For more information about SSL configuration, see "Configuring SSL." For more information about HTTPS configuration, see *Fundamentals Configuration Guide*.

The PKI domain to be referenced by the SSL policy must exist. For how to configure a PKI domain, see "Configure the PKI domain."

The configuration procedure is as follows:

1. Configure the HTTPS server:

# Configure the SSL policy for the HTTPS server to use.

```
<Device> system-view
[Device] ssl server-policy myssl
[Device-ssl-server-policy-myssl] pki-domain 1
[Device-ssl-server-policy-myssl] client-verify enable
[Device-ssl-server-policy-myssl] quit
```

2. Configure the certificate attribute group:

# Create certificate attribute group **mygroup1** and add two attribute rules. The first rule defines that the DN of the subject name includes the string **aabbcc**, and the second rule defines that the IP address of the certificate issuer is 10.0.0.1.

```
[Device] pki certificate attribute-group mygroup1
[Device-pki-cert-attribute-group-mygroup1] attribute 1 subject-name dn ctn aabbcc
[Device-pki-cert-attribute-group-mygroup1] attribute 2 issuer-name ip equ 10.0.0.1
[Device-pki-cert-attribute-group-mygroup1] quit
```

# Create certificate attribute group **mygroup2** and add two attribute rules. The first rule defines that the FQDN of the alternative subject name does not include the string of **apple**, and the second rule defines that the DN of the certificate issuer name includes the string **aabbcc**.

```
[Device] pki certificate attribute-group mygroup2
[Device-pki-cert-attribute-group-mygroup2] attribute 1 alt-subject-name fqdn nctn apple
[Device-pki-cert-attribute-group-mygroup2] attribute 2 issuer-name dn ctn aabbcc
[Device-pki-cert-attribute-group-mygroup2] quit
```

3. Configure the certificate attribute-based access control policy:

# Create the certificate attribute-based access control policy of **myacp** and add two access control rules.

```
[Device] pki certificate access-control-policy myacp
[Device-pki-cert-acp-myacp] rule 1 deny mygroup1
[Device-pki-cert-acp-myacp] rule 2 permit mygroup2
```

```
[Device-pki-cert-acp-myacp] quit
```

4. Apply the SSL server policy and certificate attribute-based access control policy to HTTPS service and enable HTTPS service:

```
Apply SSL server policy myssl to HTTPS service.
```

```
[Device] ip https ssl-server-policy myssl
```

```
Apply the certificate attribute-based access control policy of myacp to HTTPS service.
```

```
[Device] ip https certificate access-control-policy myacp
```

```
Enable HTTPS service.
```

```
[Device] ip https enable
```

## Troubleshooting PKI

### Failed to retrieve a CA certificate

#### Symptom

Failed to retrieve a CA certificate.

#### Analysis

Possible reasons include:

- The network connection is not proper. For example, the network cable might be damaged or loose.
- No trusted CA is specified.
- The URL of the registration server for certificate request is not correct or not configured.
- No authority is specified for certificate request.
- The system clock of the switch is not synchronized with that of the CA.

#### Solution

- Make sure that the network connection is physically proper.
- Check that the required commands are configured properly.
- Use the **ping** command to verify that the RA server is reachable.
- Specify the authority for certificate request.
- Synchronize the system clock of the switch with that of the CA.

### Failed to request a local certificate

#### Symptom

Failed to request a local certificate.

#### Analysis

Possible reasons include:

- The network connection is not proper. For example, the network cable might be damaged or loose.
- No CA certificate has been retrieved.
- The current key pair has been bound to a certificate.
- No trusted CA is specified.

- The URL of the registration server for certificate request is not correct or not configured.
- No authority is specified for certificate request.
- Some required parameters of the entity DN are not configured.

### Solution

- Make sure that the network connection is physically proper.
- Retrieve a CA certificate.
- Regenerate a key pair.
- Specify a trusted CA.
- Use the **ping** command to verify that the RA server is reachable.
- Specify the authority for certificate request.
- Configure the required entity DN parameters.

## Failed to retrieve CRLs

### Symptom

Failed to retrieve CRLs.

### Analysis

Possible reasons include:

- The network connection is not proper. For example, the network cable might be damaged or loose.
- No CA certificate has been retrieved before you try to retrieve CRLs.
- The IP address of LDAP server is not configured.
- The CRL distribution URL is not configured.
- The LDAP server version is wrong.

### Solution

- Make sure that the network connection is physically proper.
- Retrieve a CA certificate.
- Specify the IP address of the LDAP server.
- Specify the CRL distribution URL.
- Re-configure the LDAP version.

---

# Configuring IPsec

## Overview

IP Security (IPsec) is a security framework defined by the Internet Engineering Task Force (IETF) for securing IP communications. It is a Layer 3 virtual private network (VPN) technology that transmits data in a secure tunnel established between two endpoints.

IPsec guarantees the confidentiality, integrity, and authenticity of data and provides anti-replay service at the IP layer in an insecure network environment.

- **Confidentiality**—The sender encrypts packets before transmitting them over the Internet.
- **Data integrity**—The receiver verifies the packets received from the sender to ensure they are not tampered with during transmission.
- **Data origin authentication**—The receiver verifies the authenticity of the sender.
- **Anti-replay**—The receiver examines packets and drops outdated and duplicate packets.

IPsec delivers these benefits:

- Reduced key negotiation overheads and simplified maintenance by supporting the Internet Key Exchange (IKE) protocol. IKE provides automatic key negotiation and automatic IPsec security association (SA) setup and maintenance.
- Good compatibility. You can apply IPsec to all IP-based application systems and services without modifying them.
- Encryption on a per-packet rather than per-flow basis. Per-packet encryption allows for flexibility and greatly enhances IP security.

## IPsec implementation

IPsec comprises a set of protocols for IP data security, including Authentication Header (AH), Encapsulating Security Payload (ESP), IKE, and algorithms for authentication and encryption. AH and ESP provide security services and IKE performs key exchange.

IPsec provides two security mechanisms: authentication and encryption. The authentication mechanism allows the receiver of an IP packet to authenticate the sender and check if the packet has been tampered with. The encryption mechanism ensures data confidentiality and protects the data from being eavesdropped en route.

IPsec can use two security protocols:

- **AH (protocol 51)**—Provides data origin authentication, data integrity, and anti-replay services by adding an AH header to each IP packet. AH is suitable only for transmitting non-critical data because it cannot prevent eavesdropping, although it can prevent data tampering. AH supports authentication algorithms such as Message Digest (MD5) and Secure Hash Algorithm (SHA-1).
- **ESP (protocol 50)**—Provides data encryption as well as data origin authentication, data integrity, and anti-replay services by inserting an ESP header and an ESP trailer in IP packets. Unlike AH, ESP encrypts data before encapsulating the data to ensure data confidentiality. ESP supports encryption algorithms such as Data Encryption Standard (DES), 3DES, and Advanced Encryption Standard (AES), and authentication algorithms such as MD5 and SHA-1. The authentication function is optional to ESP.

Both AH and ESP provide authentication services, but the authentication service provided by AH is stronger. In practice, you can choose either or both security protocols. When both AH and ESP are used, an IP packet is encapsulated first by ESP and then by AH. [Figure 98](#) shows the format of IPsec packets.

## Basic concepts

### Security association

A security association is an agreement negotiated between two communicating parties called IPsec peers. It comprises a set of parameters for data protection, including security protocols, encapsulation mode, authentication and encryption algorithms, and shared keys and their lifetime. SAs can be set up manually or through IKE.

An SA is unidirectional. At least two SAs are needed to protect data flows in a bidirectional communication. If two peers want to use both AH and ESP to protect data flows between them, they construct an independent SA for each protocol.

An SA is uniquely identified by a triplet, which consists of the security parameter index (SPI), destination IP address, and security protocol identifier (AH or ESP).

An SPI is a 32-bit number for uniquely identifying an SA. It is transmitted in the AH/ESP header. A manually configured SA requires an SPI to be specified manually for it; an IKE created SA will have an SPI generated at random.

A manually configured SA never ages out. An IKE created SA has a specified period of lifetime, which comes in two types:

- Time-based lifetime, which defines how long the SA can be valid after it is created.
- Traffic-based lifetime, which defines the maximum traffic that the SA can process.

The SA becomes invalid when either of the lifetime timers expires. Before the SA expires, IKE negotiates a new SA, which takes over immediately after its creation.

### Encapsulation modes

IPsec supports the following IP packet encapsulation modes:

- **Tunnel mode**—IPsec protects the entire IP packet, including both the IP header and the payload. It uses the entire IP packet to calculate an AH or ESP header, and then encapsulates the original IP packet and the AH or ESP header with a new IP header. If you use ESP, an ESP trailer is also encapsulated. Tunnel mode is typically used for protecting gateway-to-gateway communications.
- **Transport mode**—IPsec protects only the IP payload. It uses only the IP payload to calculate the AH or ESP header, and inserts the calculated header between the original IP header and payload. If you use ESP, an ESP trailer is also encapsulated. The transport mode is typically used for protecting host-to-host or host-to-gateway communications.

[Figure 98](#) shows how the security protocols encapsulate an IP packet in different encapsulation modes.

**Figure 98 Encapsulation by security protocols in different modes**

| Mode<br>Protocol | Transport            | Tunnel                  |
|------------------|----------------------|-------------------------|
| AH               | IP AH Data           | IP AH IP Data           |
| ESP              | IP ESP Data ESP-T    | IP ESP IP Data ESP-T    |
| AH-ESP           | IP AH ESP Data ESP-T | IP AH ESP IP Data ESP-T |

## Authentication algorithms and encryption algorithms

- Authentication algorithms

IPsec uses hash algorithms to perform authentication. A hash algorithm produces a fixed-length digest for an arbitrary-length message. IPsec peers respectively calculate message digests for each packet. If the resulting digests are identical, the packet is considered intact.

IPsec supports the following hash algorithms for authentication:

- MD5, which takes as input a message of arbitrary length and produces a 128-bit message digest.
- SHA-1, which takes as input a message of a maximum length less than the 64th power of 2 in bits and produces a 160-bit message digest.

Compared with SHA-1, MD5 is faster but less secure.

- Encryption algorithms

IPsec mainly uses symmetric encryption algorithms, which encrypt and decrypt data by using the same keys. The following encryption algorithms are available for IPsec on the switch:

- Data Encryption Standard (DES), which encrypts a 64-bit plain text block with a 56-bit key. DES is the least secure but the fastest algorithm. It is sufficient for general security requirements.
- Triple DES (3DES), which encrypts plain text data with three 56-bit DES keys. The key length totals up to 168 bits. It provides moderate security strength and is slower than DES.
- Advanced Encryption Standard (AES), which encrypts plain text data with a 128-bit, 192-bit, or 256-bit key. AES provides the highest security strength and is slower than 3DES.

## IPsec SA setup modes

There are two IPsec SA setup modes:

- Manual mode. In this mode, you manually configure and maintain all SA settings. Advanced features like periodical key update are not available. However, this mode implements IPsec independently of IKE.
- ISAKMP mode. In this mode, IKE automatically negotiates and maintains IPsec SAs for IPsec.

If the number of IPsec tunnels in your network is small, use the manual mode. If the number of IPsec tunnels is large, use the ISAKMP mode.

## IPsec tunnel

An IPsec tunnel is a bidirectional channel created between two peers. An IPsec tunnel comprises one or more pairs of SAs.

## IPsec for IPv6 routing protocols

You can use IPsec to protect routing information and defend against attacks for these IPv6 routing protocols: OSPFv3, IPv6 BGP, and RIPng. The HP 5500 EI switches support using IPsec for OSPFv3, IPv6 BGP, and RIPng; the HP 5500 SI switches only support using IPsec for RIPng.

IPsec enables these IPv6 routing protocols to encapsulate outbound protocol packets and de-encapsulate inbound protocol packets with the AH or ESP protocol. If an inbound protocol packet is not IPsec protected, or fails to be de-encapsulated, for example, due to decryption or authentication failure, the routing protocol discards that packet.

You must manually configure SA parameters in an IPsec policy for IPv6 routing protocols. The IKE key exchange mechanism is applicable only to one-to-one communications. IPsec cannot implement automatic key exchange for one-to-many communications on a broadcast network, where routers must use the same SA parameters (SPI and key) to process packets for a routing protocol.

### Protocols and standards

- RFC 2401, *Security Architecture for the Internet Protocol*
- RFC 2402, *IP Authentication Header*
- RFC 2406, *IP Encapsulating Security Payload*
- RFC 4552, *Authentication/Confidentiality for OSPFv3*

## Configuring IPsec for IPv6 routing protocols

The following is the generic configuration procedure for configuring IPsec for IPv6 routing protocols:

1. Configure an IPsec proposal to specify the security protocols, authentication and encryption algorithms, and encapsulation mode.
2. Configure a manual IPsec policy to specify the keys and SPI.
3. Apply the IPsec policy to an IPv6 routing protocol.

Complete the following tasks to configure IPsec for IPv6 routing protocols:

| Task                                                 | Remarks                                                          |
|------------------------------------------------------|------------------------------------------------------------------|
| <a href="#">Configuring an IPsec proposal</a>        | Required.                                                        |
| <a href="#">Configuring an IPsec policy</a>          | Required.                                                        |
| Applying an IPsec policy to an IPv6 routing protocol | Required.<br>See <i>Layer 3—IP Routing Configuration Guide</i> . |

### Configuring an IPsec proposal

An IPsec proposal, part of an IPsec policy or an IPsec profile, defines the security parameters for IPsec SA negotiation, including the security protocol, the encryption and authentication algorithms, and the encapsulation mode.

To configure an IPsec proposal:

| Step                                                                | Command                                                                                                                                                                                                                                                                                                                                                                                                                                               | Remarks                                                                                                                                                                                                                                                                                                                                     |
|---------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                                               | <b>system-view</b>                                                                                                                                                                                                                                                                                                                                                                                                                                    | N/A                                                                                                                                                                                                                                                                                                                                         |
| 2. Create an IPsec proposal and enter its view.                     | <b>ipsec proposal</b> <i>proposal-name</i>                                                                                                                                                                                                                                                                                                                                                                                                            | By default, no IPsec proposal exists.<br><br>Optional.<br>ESP by default.                                                                                                                                                                                                                                                                   |
| 3. Specify the security protocol for the proposal.                  | <b>transform</b> { <b>ah</b>   <b>ah-esp</b>   <b>esp</b> }                                                                                                                                                                                                                                                                                                                                                                                           | Only when a security protocol is selected, can you configure security algorithms for it. For example, you can specify the ESP-specific security algorithms only when you select ESP as the security protocol. ESP supports three IP packet protection schemes: encryption only, authentication only, or both encryption and authentication. |
| 4. Specify the security algorithms.                                 | <ul style="list-style-type: none"> <li>Specify the encryption algorithm for ESP:<br/><b>esp encryption-algorithm</b> { <b>3des</b>   <b>aes</b> [ <i>key-length</i> ]   <b>des</b> }</li> <li>Specify the authentication algorithm for ESP:<br/><b>esp authentication-algorithm</b> { <b>md5</b>   <b>sha1</b> }</li> <li>Specify the authentication algorithm for AH:<br/><b>ah authentication-algorithm</b> { <b>md5</b>   <b>sha1</b> }</li> </ul> | Optional.<br><br>By default, the encryption algorithm for ESP is DES, the authentication algorithm for ESP is MD5, and the authentication algorithm for AH is MD5.                                                                                                                                                                          |
| 5. Specify the IP packet encapsulation mode for the IPsec proposal. | <b>encapsulation-mode</b> { <b>transport</b>   <b>tunnel</b> }                                                                                                                                                                                                                                                                                                                                                                                        | Optional.<br>Tunnel mode by default.<br>Transport mode applies only when the source and destination IP addresses of data flows match those of the IPsec tunnel.<br>IPsec for IPv6 routing protocols supports only the transport mode.                                                                                                       |

**NOTE:**

Changes to an IPsec proposal affect only SAs negotiated after the changes. To apply the changes to existing SAs, execute the **reset ipsec sa** command to clear the SAs so that they can be set up using the updated parameters.



## Configuring an IPsec policy

IPsec policies define which IPsec proposals should be used to protect which data flows. An IPsec policy is uniquely identified by its name and sequence number.

The switch supports only manual IPsec policies. The parameters of a manual IPsec policy are all configured manually, such as the keys and the SPIs.

### Configuration guidelines

To ensure successful SA negotiations, follow these guidelines when configuring manual IPsec policies:

- Within a certain routed network scope, the IPsec proposals used by the IPsec policies on all routers must have the same security protocols, security algorithms, and encapsulation mode. For OSPFv3, the scope can be directly connected neighbors or an OSPFv3 area. For RIPng, the scope can be directly connected neighbors or a RIPng process. For IPv6 BGP, the scope can be directly connected neighbors or a neighbor group.
- All SAs (both inbound and outbound) within the routed network scope must use the same SPI and keys.
- Configure the keys on all routers within the routed network scope in the same format. For example, if you input the keys in hexadecimal format on one router, do so across the routed network scope.

### Configuration procedure

To configure a manual IPsec policy:

| Step                                                | Command                                                                                         | Remarks                                                                                                                                                                                                                  |
|-----------------------------------------------------|-------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                               | <b>system-view</b>                                                                              | N/A                                                                                                                                                                                                                      |
| 2. Create a manual IPsec policy and enter its view. | <b>ipsec policy</b> <i>policy-name seq-number</i><br><b>manual</b>                              | By default, no IPsec policy exists.                                                                                                                                                                                      |
| 3. Assign an IPsec proposal to the IPsec policy.    | <b>proposal</b> <i>proposal-name</i>                                                            | By default, an IPsec policy references no IPsec proposal.<br>A manual IPsec policy can reference only one IPsec proposal. To change an IPsec proposal for an IPsec policy, you must remove the proposal reference first. |
| 4. Configure the local address of the tunnel        | <b>tunnel local</b> <i>ip-address</i>                                                           | Not needed for IPsec policies to be applied to IPv6 routing protocols and required for other applications.<br>Not configured by default                                                                                  |
| 5. Configure the remote address of the tunnel       | <b>tunnel remote</b> <i>ip-address</i>                                                          | Not configured by default.                                                                                                                                                                                               |
| 6. Configure the SPIs for the SAs.                  | <b>sa spi</b> { <b>inbound</b>   <b>outbound</b> } { <b>ah</b>   <b>esp</b> } <i>spi-number</i> | N/A                                                                                                                                                                                                                      |

| Step                           | Command                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | Remarks                                                                                                                                                                                                                                                                                             |
|--------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 7. Configure keys for the SAs. | <ul style="list-style-type: none"> <li>Configure an authentication key in hexadecimal for AH:<br/><b>sa authentication-hex</b> { <b>inbound</b>   <b>outbound</b> } <b>ah</b> <i>hex-key</i></li> <li>Configure an authentication key in characters for AH:<br/><b>sa string-key</b> { <b>inbound</b>   <b>outbound</b> } <b>ah</b> <i>string-key</i></li> <li>Configure a key in characters for ESP:<br/><b>sa string-key</b> { <b>inbound</b>   <b>outbound</b> } <b>esp</b> <i>string-key</i></li> <li>Configure an authentication key in hexadecimal for ESP:<br/><b>sa authentication-hex</b> { <b>inbound</b>   <b>outbound</b> } <b>esp</b> <i>hex-key</i></li> <li>Configure an encryption key in hexadecimal for ESP:<br/><b>sa encryption-hex</b> { <b>inbound</b>   <b>outbound</b> } <b>esp</b> <i>hex-key</i></li> </ul> | <p>Configure a key for AH, ESP, or both.</p> <p>If you configure a key in characters for ESP, the router automatically generates an authentication key and an encryption key for ESP.</p> <p>If you configure a key in two modes: string and hexadecimal, the last configured one takes effect.</p> |

## Displaying and maintaining IPsec

| Task                                | Command                                                                                                                                                                                | Remarks                |
|-------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|
| Display IPsec policy information.   | <b>display ipsec policy</b> [ <b>brief</b>   <b>name</b> <i>policy-name</i> [ <i>seq-number</i> ] ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] | Available in any view  |
| Display IPsec proposal information. | <b>display ipsec proposal</b> [ <i>proposal-name</i> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                                              | Available in any view  |
| Display IPsec SA information.       | <b>display ipsec sa</b> [ <b>brief</b>   <b>policy</b> <i>policy-name</i> [ <i>seq-number</i> ] ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]   | Available in any view  |
| Display IPsec packet statistics.    | <b>display ipsec statistics</b> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                                                                     | Available in any view  |
| Clear SAs.                          | <b>reset ipsec sa</b> [ <b>policy</b> <i>policy-name</i> [ <i>seq-number</i> ] ]                                                                                                       | Available in user view |
| Clear IPsec statistics.             | <b>reset ipsec statistics</b>                                                                                                                                                          | Available in user view |

## IPsec for RIPng configuration example

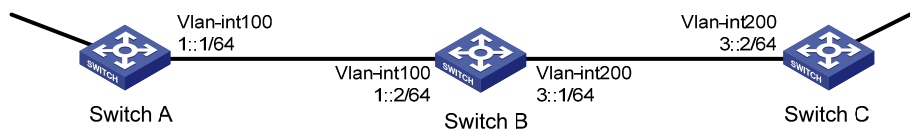
The IPsec configuration procedures for protecting OSPFv3 and IPv6 BGP are similar. For more information about RIPng, OSPFv3, and IPv6 BGP, see *Layer 3—IP Routing Configuration Guide*. Only the HP 5500 EI switches support IPsec for OSPFv3 and IPv6 BGP.

## Network requirements

As shown in [Figure 99](#), Switch A, Switch B, and Switch C are connected. They learn IPv6 routing information through RIPng.

Configure IPsec for RIPng so that RIPng packets exchanged between the switches are transmitted through an IPsec tunnel. Configure IPsec to use the security protocol ESP, the encryption algorithm DES, and the authentication algorithm SHA1-HMAC-96.

**Figure 99 Network diagram**



## Configuration considerations

To meet the requirements, perform the following configuration tasks:

- Configure basic RIPng parameters.
- Configure a manual IPsec policy.
- Apply the IPsec policy to a RIPng process to protect RIPng packets in this process or to an interface to protect RIPng packets traveling through the interface.

## Configuration procedure

### 1. Configure Switch A:

# Assign an IPv6 address to each interface. (Details not shown.)

# Create a RIPng process and enable it on VLAN-interface 100.

```
<SwitchA> system-view
[SwitchA] ripng 1
[SwitchA-ripng-1] quit
[SwitchA] interface vlan-interface 100
[SwitchA-Vlan-interface100] ripng 1 enable
[SwitchA-Vlan-interface100] quit
```

# Create an IPsec proposal named **tran1**, and set the encapsulation mode to transport mode, the security protocol to ESP, the encryption algorithm to DES, and authentication algorithm to SHA1-HMAC-96.

```
[SwitchA] ipsec proposal tran1
[SwitchA-ipsec-proposal-tran1] encapsulation-mode transport
[SwitchA-ipsec-proposal-tran1] transform esp
[SwitchA-ipsec-proposal-tran1] esp encryption-algorithm des
[SwitchA-ipsec-proposal-tran1] esp authentication-algorithm sha1
[SwitchA-ipsec-proposal-tran1] quit
```

# Create an IPsec policy named **policy001**, specify the manual mode for it, and set the SPIs of the inbound and outbound SAs to 123456, and the keys for the inbound and outbound SAs using ESP to abcdefg.

```
[SwitchA] ipsec policy policy001 10 manual
[SwitchA-ipsec-policy-manual-policy001-10] proposal tran1
[SwitchA-ipsec-policy-manual-policy001-10] sa spi outbound esp 123456
[SwitchA-ipsec-policy-manual-policy001-10] sa spi inbound esp 123456
[SwitchA-ipsec-policy-manual-policy001-10] sa string-key outbound esp abcdefg
```

```
[SwitchA-ipsec-policy-manual-policy001-10] sa string-key inbound esp abcdefg
[SwitchA-ipsec-policy-manual-policy001-10] quit
Apply IPsec policy policy001 to the RIPng process.
[SwitchA] ripng 1
[SwitchA-ripng-1] enable ipsec-policy policy001
[SwitchA-ripng-1] quit
```

## 2. Configure Switch B

# Assign an IPv6 address to each interface. (Details not shown.)

# Create a RIPng process and enable it on VLAN-interface 100 and VLAN-interface 200.

```
<SwitchB> system-view
[SwitchB] ripng 1
[SwitchB-ripng-1] quit
[SwitchB] interface vlan-interface 200
[SwitchB-Vlan-interface200] ripng 1 enable
[SwitchB-Vlan-interface200] quit
[SwitchB] interface vlan-interface 100
[SwitchB-Vlan-interface100] ripng 1 enable
[SwitchB-Vlan-interface100] quit
```

# Create an IPsec proposal named **tran1**, and set the encapsulation mode to transport mode, the security protocol to ESP, the encryption algorithm to DES, and authentication algorithm to SHA1-HMAC-96.

```
[SwitchB] ipsec proposal tran1
[SwitchB-ipsec-proposal-tran1] encapsulation-mode transport
[SwitchB-ipsec-proposal-tran1] transform esp
[SwitchB-ipsec-proposal-tran1] esp encryption-algorithm des
[SwitchB-ipsec-proposal-tran1] esp authentication-algorithm sha1
[SwitchB-ipsec-proposal-tran1] quit
```

# Create an IPsec policy named **policy001**, specify the manual mode for it, and configure the SPIs of the inbound and outbound SAs to 123456, and the keys for the inbound and outbound SAs using ESP to abcdefg.

```
[SwitchB] ipsec policy policy001 10 manual
[SwitchB-ipsec-policy-manual-policy001-10] proposal tran1
[SwitchB-ipsec-policy-manual-policy001-10] sa spi outbound esp 123456
[SwitchB-ipsec-policy-manual-policy001-10] sa spi inbound esp 123456
[SwitchB-ipsec-policy-manual-policy001-10] sa string-key outbound esp abcdefg
[SwitchB-ipsec-policy-manual-policy001-10] sa string-key inbound esp abcdefg
[SwitchB-ipsec-policy-manual-policy001-10] quit
```

# Apply IPsec policy **policy001** to the RIPng process.

```
[SwitchB] ripng 1
[SwitchB-ripng-1] enable ipsec-policy policy001
[SwitchB-ripng-1] quit
```

## 3. Configure Switch C

# Assign an IPv6 address to each interface. (Details not shown.)

# Create a RIPng process and enable it on VLAN-interface 200.

```
<SwitchC> system-view
[SwitchC] ripng 1
```

```
[SwitchC-ripng-1] quit
[SwitchC] interface vlan-interface 200
[SwitchC-Vlan-interface200] ripng 1 enable
[SwitchC-Vlan-interface200] quit
```

# Create an IPsec proposal named **tran1**, and set the encapsulation mode to transport mode, the security protocol to ESP, the encryption algorithm to DES, and authentication algorithm to SHA1-HMAC-96.

```
[SwitchC] ipsec proposal tran1
[SwitchC-ipsec-proposal-tran1] encapsulation-mode transport
[SwitchC-ipsec-proposal-tran1] transform esp
[SwitchC-ipsec-proposal-tran1] esp encryption-algorithm des
[SwitchC-ipsec-proposal-tran1] esp authentication-algorithm sha1
[SwitchC-ipsec-proposal-tran1] quit
```

# Create an IPsec policy named **policy001**, specify the manual mode for it, and configure the SPIs of the inbound and outbound SAs to 123456, and the keys for the inbound and outbound SAs using ESP to abcdefg.

```
[SwitchC] ipsec policy policy001 10 manual
[SwitchC-ipsec-policy-manual-policy001-10] proposal tran1
[SwitchC-ipsec-policy-manual-policy001-10] sa spi outbound esp 123456
[SwitchC-ipsec-policy-manual-policy001-10] sa spi inbound esp 123456
[SwitchC-ipsec-policy-manual-policy001-10] sa string-key outbound esp abcdefg
[SwitchC-ipsec-policy-manual-policy001-10] sa string-key inbound esp abcdefg
[SwitchC-ipsec-policy-manual-policy001-10] quit
```

# Apply IPsec policy **policy001** to the RIPng process.

```
[SwitchC] ripng 1
[SwitchC-ripng-1] enable ipsec-policy policy001
[SwitchC-ripng-1] quit
```

#### 4. Verify the configuration

After the configuration, Switch A, Switch B, and Switch C learns IPv6 routing information through RIPng. SAs are set up successfully, and the IPsec tunnel between two peers is up for protecting the RIPng packets.

Using the **display ripng** command on Switch A, you will see the running status and configuration information of the specified RIPng process. The output shows that IPsec policy policy001 is applied to this process successfully.

```
<SwitchA> display ripng 1
 RIPng process : 1
 Preference : 100
 Checkzero : Enabled
 Default Cost : 0
 Maximum number of balanced paths : 8
 Update time : 30 sec(s) Timeout time : 180 sec(s)
 Suppress time : 120 sec(s) Garbage-Collect time : 120 sec(s)
 Number of periodic updates sent : 186
 Number of trigger updates sent : 1
 IPsec policy name: policy001, SPI: 123456
```

Using the **display ipsec sa** command on Switch A, you will see the information about the inbound and outbound SAs.

```
<SwitchA> display ipsec sa
=====
Protocol: RIPng
=====

IPsec policy name: "policy001"
sequence number: 10
mode: manual

connection id: 1
encapsulation mode: transport
perfect forward secrecy:
tunnel:
flow:

[inbound ESP SAs]
spi: 123456 (0x3039)
proposal: ESP-ENCRYPT-DES ESP-AUTH-SHA1
No duration limit for this sa

[outbound ESP SAs]
spi: 123456 (0x3039)
proposal: ESP-ENCRYPT-DES ESP-AUTH-SHA1
No duration limit for this sa
```

Similarly, you can view the information on Switch B and Switch C. (Details not shown.)

# Configuring SSH2.0

## Overview

### Introduction to SSH2.0

Secure Shell (SSH) offers an approach to logging in to a remote device securely. Using encryption and strong authentication, SSH protects devices against attacks such as IP spoofing and plain text password interception.

The switch can not only work as an SSH server to support connections with SSH clients, but also work as an SSH client to allow users to establish SSH connections with a remote device acting as the SSH server.

Unless otherwise noted, SSH in this document refers to SSH2.0.

---

**NOTE:**

When acting as an SSH server, the switch supports SSH2.0 and SSH1. When acting as an SSH client, the switch supports SSH2.0 only.

---

## SSH operation

To establish an SSH connection and communicate with each other through the connection, an SSH client and the SSH server go through the stages listed in [Table 14](#).

**Table 14 Stages in session establishment and interaction between an SSH client and the server**

| Stages                        | Description                                                                                                                                                                      |
|-------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Version negotiation           | SSH1 and SSH2.0 are supported. The two parties negotiate a version to use.                                                                                                       |
| Key and algorithm negotiation | SSH supports multiple algorithms. The two parties negotiate algorithms for communication, and use the DH key exchange algorithm to generate the same session key and session ID. |
| Authentication                | The SSH server authenticates the client in response to the client's authentication request.                                                                                      |
| Session request               | After passing authentication, the client sends a session request to the server.                                                                                                  |
| Interaction                   | After the server grants the request, the client and the server start to communicate with each other.                                                                             |

### Version negotiation

1. The server opens port 22 to listen to connection requests from clients.
2. The client sends a TCP connection request to the server.
3. After the TCP connection is established, the server sends a packet that carries a version information string to the client. The version information string is in the format SSH-<primary protocol version number>.<secondary protocol version number>-<software version number>. The primary and

secondary protocol version numbers constitute the protocol version number. The software version number is used for debugging.

4. After receiving the packet, the client resolves the packet and compares the server's protocol version number with that of its own. If the server's protocol version is lower and supportable, the client uses the protocol version of the server; otherwise, the client uses its own protocol version. In either case, the client sends a packet to the server to notify the server of the protocol version that it decides to use.
5. The server compares the version number carried in the packet with that of its own. If the server supports the version, the negotiation succeeds and the server and the client proceed with key and algorithm negotiation. Otherwise, the negotiation fails, and the server breaks the TCP connection.

---

**NOTE:**

All the packets involved in the preceding steps are transferred in plain text.

---

## Key and algorithm negotiation



**IMPORTANT:**

Before the key and algorithm negotiation, the server must have already generated a DSA or RSA key pair, which is used in generating the session key and session ID, and by the client to authenticate the identity of the server. For more information about DSA and RSA key pairs, see "[Managing public keys.](#)"

---

The server and the client send algorithm negotiation packets to each other, notifying the peer of the supported public key algorithms, encryption algorithms, Message Authentication Code (MAC) algorithms, and compression algorithms.

Based on the received algorithm negotiation packets, the server and the client figure out the algorithms to be used. If the negotiation of any type of algorithm fails, the algorithm negotiation fails and the server tears down the connection with the client.

The server and the client use the DH key exchange algorithm and parameters such as the host key pair to generate the session key and session ID, and the client authenticates the identity of the server.

Through the steps, the server and the client get the same session key and session ID. The session key will be used to encrypt and decrypt data exchanged between the server and client later. The session ID will be used to identify the session established between the server and client and will be used in the authentication stage.

## Authentication

SSH supports the following authentication methods:

- **Password authentication**—The SSH server uses AAA for authentication of the client. During password authentication, the SSH client encrypts its username and password, encapsulates them into a password authentication request, and sends the request to the server. After receiving the request, the SSH server decrypts the username and password, checks the validity of the username and password locally or by a remote AAA server, and then informs the client of the authentication result. If the remote AAA server requires the user for a password re-authentication, it carries a prompt in the authentication response sent to the client. The prompt is transparently transmitted to the client, and displayed on the client to notify the user to enter a specified password. After the user enters the correct password and passes validity check on the remote AAA server, the server returns an authentication success message to the client.
- **Publickey authentication**—The server authenticates the client by the digital signature. During publickey authentication, the client sends the server a publickey authentication request that contains



its username, public key, and publickey algorithm information. The server checks whether the public key is valid. If the public key is invalid, the authentication fails. Otherwise, the server authenticates the client by the digital signature. Finally, the server sends a message to the client to inform it of the authentication result. The switch supports using the publickey algorithms RSA and DSA for digital signature.

An SSH2.0 server might require the client to pass both password authentication and publickey authentication or either of them. However, if the client is running SSH1, the client only needs to pass either authentication, regardless of the requirement of the server.

The following gives the steps of the authentication stage:

1. The client sends the server an authentication request that includes the username, the authentication method, and the information related to the authentication method (for example, the password in the case of password authentication).
2. The server authenticates the client. If the authentication fails, the server sends the client a message to inform the client of the failure and the methods available for re-authentication.
3. The client selects a method from the list to initiate another authentication.
4. The preceding process repeats until the authentication succeeds or the number of failed authentication attempts exceeds the maximum of authentication attempts. In the latter case, the server tears the session down.

---

**NOTE:**

Only clients running SSH2.0 or a later version support password re-authentication that is initiated by the switch acting as the SSH server.

---

## Session request

After passing authentication, the client sends a session request to the server, and the server listens to and processes the request from the client. If the server successfully processes the request, the server sends an SSH\_MSG\_SUCCESS packet to the client and goes on to the interaction stage with the client. Otherwise, the server sends an SSH\_MSG\_FAILURE packet to the client to indicate that the processing has failed or it cannot resolve the request.

## Interaction

In this stage, the server and the client exchanges data as follows:

1. The client encrypts and sends the command to be executed to the server.
2. The server decrypts and executes the command, and then encrypts and sends the result to the client.
3. The client decrypts and displays the result on the terminal.

In the interaction stage, you can execute commands from the client by pasting the commands in text format (the text must be within 2000 bytes). The commands must be available in the same view. Otherwise, the server might not be able to perform the commands correctly.

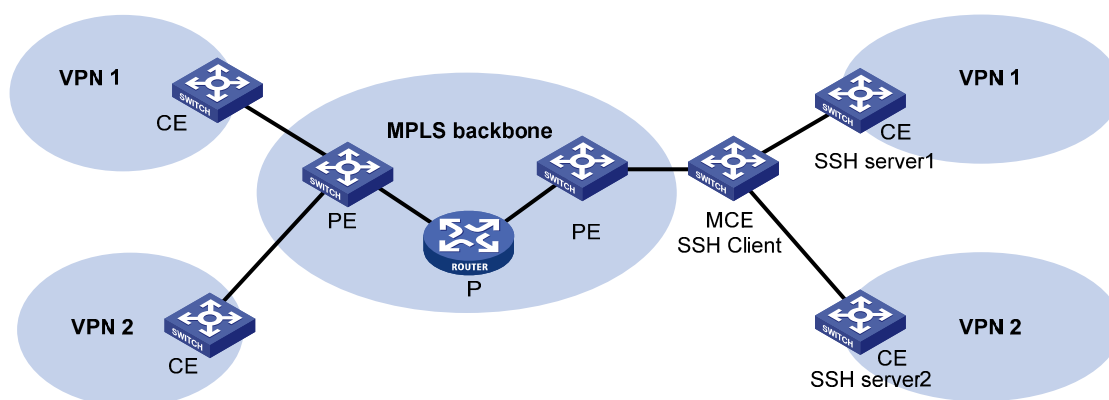
If the command text exceeds 2000 bytes, you can execute the commands by saving the text as a configuration file, uploading the configuration file to the server through Secure FTP (SFTP), and then using the configuration file to restart the server.

# SSH connection across VPNs (available only on the HP 5500 EI)

With this function, you can configure the device as an SSH client to establish connections with SSH servers in different VPNs.

As shown in [Figure 100](#), the hosts in VPN 1 and VPN 2 access the MPLS backbone through PEs, with the services of the two VPNs isolated. After an HP 5500 EI switch that acts as an MCE device is enabled with the SSH client function, it can establish SSH connections with CEs in different VPNs that are enabled with the SSH server function to implement secure access to the CEs and secure transfer of log file.

**Figure 100 Network diagram**



For more information about MCE, see *Layer 3—IP Routing Configuration Guide*.

## Configuring the switch as an SSH server

### SSH server configuration task list

| Task                                                                      | Remarks                                                                                    |
|---------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|
| <a href="#">Generating DSA or RSA key pairs</a>                           | Required                                                                                   |
| <a href="#">Enabling the SSH server function</a>                          | Required                                                                                   |
| <a href="#">Configuring the user interfaces for SSH clients</a>           | Required                                                                                   |
| <a href="#">Configuring a client public key</a>                           | Required for publickey authentication users and optional for password authentication users |
| <a href="#">Configuring an SSH user</a>                                   | Optional                                                                                   |
| <a href="#">Setting the SSH management parameters</a>                     | Optional                                                                                   |
| <a href="#">Setting the DSCP value for packets sent by the SSH server</a> | Optional                                                                                   |

### Generating DSA or RSA key pairs

In the key and algorithm negotiation stage, the DSA or RSA key pairs are used to generate the session key and session ID and for the client to authenticate the server.

To support SSH clients that use different types of key pairs, generate both DSA and RSA key pairs on the SSH server.

## Generating procedure

To generate DSA or RSA key pairs on the SSH server:

| Step                              | Command                                      | Remarks                                          |
|-----------------------------------|----------------------------------------------|--------------------------------------------------|
| 1. Enter system view.             | <b>system-view</b>                           | N/A                                              |
| 2. Generate DSA or RSA key pairs. | <b>public-key local create { dsa   rsa }</b> | By default, neither DSA nor RSA key pairs exist. |

## Commands for generating DSA or RSA key pairs

The **public-key local create rsa** command generates a server RSA key pair and a host RSA key pair. Each of the key pairs consists of a public key and a private key. The public key in the server key pair of the SSH server is used in SSH1 to encrypt the session key for secure transmission of the key. As SSH2.0 uses the DH algorithm to generate the session key on the SSH server and client, no session key transmission is required in SSH2.0 and the server key pair is not used.

The length of the modulus of RSA server keys and host keys must be in the range of 512 to 2048 bits. Some SSH2.0 clients require that the length of the key modulus be at least 768 bits on the SSH server side.

The **public-key local create dsa** command generates only the host key pair. SSH1 does not support the DSA algorithm.

The length of the modulus of DSA host keys must be in the range of 512 to 2048 bits. Some SSH2.0 clients require that the length of the key modulus be at least 768 bits on the SSH server side.

For more information about the **public-key local create** command, see *Security Command Reference*.

## Enabling the SSH server function

| Step                               | Command                  | Remarks             |
|------------------------------------|--------------------------|---------------------|
| 1. Enter system view.              | <b>system-view</b>       | N/A                 |
| 2. Enable the SSH server function. | <b>ssh server enable</b> | Disabled by default |

## Configuring the user interfaces for SSH clients

### Configuration guidelines

An SSH client accesses the switch through a VTY user interface. You must configure the user interfaces for SSH clients to allow SSH login. The configuration takes effect only for clients that log in after the configuration.

If you configure a user interface to support SSH, be sure to configure the corresponding authentication mode with the **authentication-mode scheme** command.

For a user interface configured to support SSH, you cannot change the authentication mode. To change the authentication mode, undo the SSH support configuration first.

## Configuration procedure

To configure the protocols for a user interface to support:

| Step                                                         | Command                                                             | Remarks                                                  |
|--------------------------------------------------------------|---------------------------------------------------------------------|----------------------------------------------------------|
| 1. Enter system view.                                        | <b>system-view</b>                                                  | N/A                                                      |
| 2. Enter user interface view of one or more user interfaces. | <b>user-interface vty</b> <i>number</i><br>[ <i>ending-number</i> ] | N/A                                                      |
| 3. Set the login authentication mode to <b>scheme</b> .      | <b>authentication-mode scheme</b>                                   | By default, the authentication mode is <b>password</b> . |
| 4. Configure the user interface(s) to support SSH login.     | <b>protocol inbound</b> { <b>all</b>   <b>ssh</b> }                 | Optional.<br>All protocols are supported by default.     |

For more information about the **authentication-mode** and **protocol inbound** commands, see *Fundamentals Command Reference*.

## Configuring a client public key

This configuration task is only necessary for SSH users using publickey authentication.

To allow an SSH user to pass publickey authentication and log in to the server, you must configure the client's DSA or RSA host public key on the server, and configure the client to use the corresponding host private key, so that the server uses the digital signature to authenticate the client.

You can manually configure the public key of an SSH client on the server, or import it from the public key file:

- **Configure it manually**—You can type or copy the public key to the SSH server. The public key must have not been converted and be in the Distinguished Encoding Rules (DER) encoding format.
- **Import it from the public key file**—During the import process, the server will automatically convert the public key in the public key file to a string in Public Key Cryptography Standards (PKCS) format, and save it locally. Before importing the public key, you must upload the public key file (in binary) to the server through FTP or TFTP.

### NOTE:

HP recommends you to configure a client public key by importing it from a public key file.

## Configuring a client public key manually

| Step                                                                  | Command                                  | Remarks                                                                            |
|-----------------------------------------------------------------------|------------------------------------------|------------------------------------------------------------------------------------|
| 1. Enter system view.                                                 | <b>system-view</b>                       | N/A                                                                                |
| 2. Enter public key view.                                             | <b>public-key peer</b> <i>keyname</i>    | N/A                                                                                |
| 3. Enter public key code view.                                        | <b>public-key-code begin</b>             | N/A                                                                                |
| 4. Configure a client's host public key.                              | Enter the content of the host public key | Spaces and carriage returns are allowed between characters.                        |
| 5. Return to public key view and save the configured host public key. | <b>public-key-code end</b>               | When you exit public key code view, the system automatically saves the public key. |

| Step                      | Command                    | Remarks |
|---------------------------|----------------------------|---------|
| 6. Return to system view. | <b>peer-public-key end</b> | N/A     |

## Importing a client public key from a public key file

| Step                                             | Command                                                                    |
|--------------------------------------------------|----------------------------------------------------------------------------|
| 1. Enter system view.                            | <b>system-view</b>                                                         |
| 2. Import the public key from a public key file. | <b>public-key peer</b> <i>keyname</i> <b>import sshkey</b> <i>filename</i> |

For more information about client public key configuration, see "[Managing public keys](#)."

## Configuring an SSH user

To configure an SSH user that uses publickey authentication, you must perform the procedure in this section.

To configure an SSH user that uses password authentication, whether together with publickey authentication or not, you must configure a local user account by using the **local-user** command in "[Configuring AAA](#)" for local authentication, or configure an SSH user account on an authentication server, for example, a RADIUS server, for remote authentication.

For password-only SSH users, you do not need to perform the procedure in this section to configure them unless you want to use the **display ssh user-information** command to display all SSH users, including the password-only SSH users, for centralized management.

### Configuration guidelines

When you perform the procedure in this section to configure an SSH user, follow these guidelines:

You can set the service type to Stelnet, SFTP, and SCP (Secure copy). For more information about Stelnet, see "[Overview](#)." For more information about SFTP, see "[Configuring SFTP](#)." For more information about SCP, see "[Configuring SCP](#)."

- You can enable one of the following authentication modes for the SSH user:
  - **Password**—The user must pass password authentication.
  - **Publickey authentication**—The user must pass publickey authentication.
  - **Password-publickey authentication**—As an SSH2.0 user, the user must pass both password and publickey authentication. As an SSH1 user, the user must pass either password or publickey authentication.
  - **Any**—The user can use either password authentication or publickey authentication.
- If publickey authentication, whether with password authentication or not, is used, the command level accessible to the user is set by the **user privilege level** command on the user interface. If only password authentication is used, the command level accessible to the user is authorized by AAA.
- SSH1 does not support SCP and SFTP. For an SSH1 client, you must set the service type to **stelnet** or **all**.
- For an SCP or SFTP user, the working folder depends on the authentication method:
  - If only password authentication is used, the working folder is authorized by AAA.
  - If publickey authentication, whether with password authentication or not, is used, the working folder is set by using the **ssh user** command.

- If you change the authentication mode or public key for an SSH user that has been logged in, the change can take effect only at the next login of the user.

## Configuration procedure

To configure an SSH user and specify the service type and authentication method:

| Step                                                                           | Command                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | Remarks             |
|--------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| 1. Enter system view.                                                          | <b>system-view</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | N/A                 |
| 2. Create an SSH user, and specify the service type and authentication method. | <ul style="list-style-type: none"> <li>• For Stelnet users:<br/> <b>ssh user <i>username</i> service-type stelnet authentication-type { password   { any   password-publickey   publickey } assign publickey <i>keyname</i> }</b> </li> <li>• For all users, SCP or SFTP users:<br/> <b>ssh user <i>username</i> service-type { all   scp   sftp } authentication-type { password   { any   password-publickey   publickey } assign publickey <i>keyname</i> work-directory <i>directory-name</i> }</b> </li> </ul> | Use either command. |

## Setting the SSH management parameters

SSH management includes:

- Enabling the SSH server to be compatible with SSH1 client
- Setting the RSA server key pair update interval, applicable to users using SSH1 client
- Setting the SSH user authentication timeout period
- Setting the maximum number of SSH authentication attempts

Setting these parameters can help avoid malicious guessing at and cracking of the keys and usernames, securing your SSH connections.



### IMPORTANT:

Authentication fails if the number of authentication attempts (including both publickey and password authentication) exceeds that specified in the **ssh server authentication-retries** command.

To set the SSH management parameters:

| Step                                              | Command                                       | Remarks                                                        |
|---------------------------------------------------|-----------------------------------------------|----------------------------------------------------------------|
| 1. Enter system view.                             | <b>system-view</b>                            | N/A                                                            |
| 2. Enable the SSH server to support SSH1 clients. | <b>ssh server compatible-ssh1x [ enable ]</b> | Optional.<br>By default, the SSH server supports SSH1 clients. |

| Step                                                      | Command                                                        | Remarks                                                                                 |
|-----------------------------------------------------------|----------------------------------------------------------------|-----------------------------------------------------------------------------------------|
| 3. Set the RSA server key pair update interval.           | <b>ssh server rekey-interval</b> <i>hours</i>                  | Optional.<br>By default, the interval is 0, and the RSA server key pair is not updated. |
| 4. Set the SSH user authentication timeout period.        | <b>ssh server authentication-timeout</b> <i>time-out-value</i> | Optional.<br>60 seconds by default.                                                     |
| 5. Set the maximum number of SSH authentication attempts. | <b>ssh server authentication-retries</b> <i>times</i>          | Optional.<br>3 by default.                                                              |

## Setting the DSCP value for packets sent by the SSH server

| Step                                                      | Command                                                                                                                                                                                                                                                                           | Remarks                                                                                                                               |
|-----------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                                     | <b>system-view</b>                                                                                                                                                                                                                                                                | N/A                                                                                                                                   |
| 2. Set the DSCP value for packets sent by the SSH server. | <ul style="list-style-type: none"> <li>Set the DSCP value for packets sent by the IPv4 SSH server:<br/><b>ssh server dscp</b> <i>dscp-value</i></li> <li>Set the DSCP value for packets sent by the IPv6 SSH server:<br/><b>ssh server ipv6 dscp</b> <i>dscp-value</i></li> </ul> | Optional.<br>By default, the DSCP value is 16 in packets sent by the IPv4 SSH server and is 0 in packets sent by the IPv6 SSH server. |

## Configuring the switch as an SSH client

### SSH client configuration task list

| Task                                                                        | Remarks  |
|-----------------------------------------------------------------------------|----------|
| <a href="#">Specifying a source IP address/interface for the SSH client</a> | Optional |
| <a href="#">Configuring whether first-time authentication is supported</a>  | Optional |
| <a href="#">Establishing a connection between the SSH client and server</a> | Required |
| <a href="#">Setting the DSCP value for packets sent by the SSH client</a>   | Optional |

### Specifying a source IP address/interface for the SSH client

This configuration task allows you to specify a source IP address or interface for the client to access the SSH server, improving service manageability.

To specify a source IP address or interface for the client:

| Step                  | Command            | Remarks |
|-----------------------|--------------------|---------|
| 1. Enter system view. | <b>system-view</b> | N/A     |

| Step                                                            | Command                                                                                                                                                                                                                                                                                                                                                                                                                                         | Remarks                                                                                                                                                                        |
|-----------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2. Specify a source IP address or interface for the SSH client. | <ul style="list-style-type: none"> <li>Specify a source IPv4 address or interface for the SSH client:<br/><b>ssh client source</b> { <b>ip</b> <i>ip-address</i>   <b>interface</b> <i>interface-type interface-number</i> }</li> <li>Specify a source IPv6 address or interface for the SSH client:<br/><b>ssh client ipv6 source</b> { <b>ipv6</b> <i>ipv6-address</i>   <b>interface</b> <i>interface-type interface-number</i> }</li> </ul> | <p>Select either approach.</p> <p>By default, an SSH client uses the IP address of the outbound interface defined by the route to the SSH server to access the SSH server.</p> |

## Configuring whether first-time authentication is supported

When the switch acts as an SSH client and connects to the SSH server, you can configure whether the switch supports first-time authentication.

- With first-time authentication, when an SSH client not configured with the server host public key accesses the server for the first time, the user can continue accessing the server, and save the host public key on the client. When accessing the server again, the client will use the saved server host public key to authenticate the server.
- Without first-time authentication, a client not configured with the server host public key will refuse to access the server. To enable the client to access the server, you must configure the server host public key and specify the public key name for authentication on the client in advance.

### Enabling the switch to support first-time authentication

| Step                                                       | Command                                        | Remarks                                                                      |
|------------------------------------------------------------|------------------------------------------------|------------------------------------------------------------------------------|
| 1. Enter system view.                                      | <b>system-view</b>                             | N/A                                                                          |
| 2. Enable the switch to support first-time authentication. | <b>ssh client first-time</b> [ <b>enable</b> ] | Optional.<br>By default, first-time authentication is supported on a client. |

### Disabling first-time authentication

For successful authentication of an SSH client not supporting first-time authentication, the server host public key must be configured on the client and the public key name must be specified.

To disable first-time authentication:

| Step                                               | Command                                                                                  | Remarks                                                                                                                                 |
|----------------------------------------------------|------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                              | <b>system-view</b>                                                                       | N/A                                                                                                                                     |
| 2. Disable first-time authentication support.      | <b>undo ssh client first-time</b>                                                        | By default, first-time authentication is supported on a client.                                                                         |
| 3. Configure the server host public key.           | See " <a href="#">Configuring a client public key</a> "                                  | The method for configuring the server host public key on the client is similar to that for configuring client public key on the server. |
| 4. Specify the host public key name of the server. | <b>ssh client authentication server</b><br><b>server assign publickey</b> <i>keyname</i> | N/A                                                                                                                                     |



## Establishing a connection between the SSH client and server

| Task                                                                                                                                                                                               | Command                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | Remarks                                                                                                             |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------|
| Establish a connection between the SSH client and the server, and specify the public key algorithm, preferred encryption algorithm, preferred HMAC algorithm and preferred key exchange algorithm. | <ul style="list-style-type: none"> <li>For an IPv4 server:<br/> <code>ssh2 server [ port-number ] [ vpn-instance vpn-instance-name ] [ identity-key { dsa   rsa }   prefer-ctos-cipher { 3des   aes128   des }   prefer-ctos-hmac { md5   md5-96   sha1   sha1-96 }   prefer-kex { dh-group-exchange   dh-group1   dh-group14 }   prefer-stoc-cipher { 3des   aes128   des }   prefer-stoc-hmac { md5   md5-96   sha1   sha1-96 } ] *</code> </li> <li>For an IPv6 server:<br/> <code>ssh2 ipv6 server [ port-number ] [ vpn-instance vpn-instance-name ] [ identity-key { dsa   rsa }   prefer-ctos-cipher { 3des   aes128   des }   prefer-ctos-hmac { md5   md5-96   sha1   sha1-96 }   prefer-kex { dh-group-exchange   dh-group1   dh-group14 }   prefer-stoc-cipher { 3des   aes128   des }   prefer-stoc-hmac { md5   md5-96   sha1   sha1-96 } ] *</code> </li> </ul> | <p>Use either command in user view.</p> <p>Only the HP 5500 EI switches support the <b>vpn-instance</b> option.</p> |

## Setting the DSCP value for packets sent by the SSH client

| Step                                                      | Command                                                                                                                                                                                                                                                                             | Remarks                                                                                                                                          |
|-----------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                                     | <code>system-view</code>                                                                                                                                                                                                                                                            | N/A                                                                                                                                              |
| 2. Set the DSCP value for packets sent by the SSH client. | <ul style="list-style-type: none"> <li>Set the DSCP value for packets sent by the IPv4 SSH client:<br/> <code>ssh client dscp dscp-value</code> </li> <li>Set the DSCP value for packets sent by the IPv6 SSH client:<br/> <code>ssh client ipv6 dscp dscp-value</code> </li> </ul> | <p>Optional.</p> <p>By default, the DSCP value is 16 in packets sent by the IPv4 SSH client and is 0 in packets sent by the IPv6 SSH client.</p> |

## Displaying and maintaining SSH

| Task                                                                     | Command                                                                                        | Remarks               |
|--------------------------------------------------------------------------|------------------------------------------------------------------------------------------------|-----------------------|
| Display the source IP address or interface set for the SFTP client.      | <code>display sftp client source [   { begin   exclude   include } regular-expression ]</code> | Available in any view |
| Display the source IP address or interface information on an SSH client. | <code>display ssh client source [   { begin   exclude   include } regular-expression ]</code>  | Available in any view |

| Task                                                                                  | Command                                                                                                                                                                | Remarks               |
|---------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Display SSH server status information or session information on an SSH server.        | <b>display ssh server</b> { <b>status</b>   <b>session</b> } [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                        | Available in any view |
| Display the mappings between SSH servers and their host public keys on an SSH client. | <b>display ssh server-info</b> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                                                      | Available in any view |
| Display information about SSH users on an SSH server.                                 | <b>display ssh user-information</b> [ <i>username</i> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                             | Available in any view |
| Display the public keys of the local key pairs.                                       | <b>display public-key local</b> { <b>dsa</b>   <b>rsa</b> } <b>public</b> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]           | Available in any view |
| Display the public keys of the SSH peers.                                             | <b>display public-key peer</b> [ <b>brief</b>   <b>name</b> <i>publickey-name</i> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] | Available in any view |

For more information about the **display public-key local** and **display public-key peer** commands, see *Security Command Reference*.

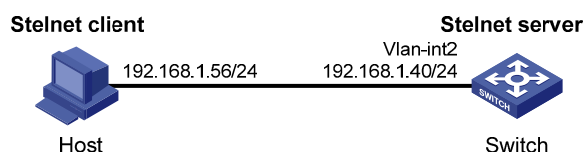
## SSH server configuration examples

### When the switch acts as a server for password authentication

#### Network requirements

As shown in [Figure 101](#), a host (the SSH client) and a switch (the SSH server) are directly connected. Configure an SSH user on the switch so that the host can securely log in to the switch after passing password authentication. Configure a username and password for the user on the switch.

**Figure 101 Network diagram**



#### Configuration procedure

1. Configure the SSH server:

```
Generate the RSA key pairs.
```

```
<Switch> system-view
```

```
[Switch] public-key local create rsa
```

```
The range of public key size is (512 ~ 2048).
```

```
NOTES: If the key modulus is greater than 512,
```

```
It will take a few minutes.
```

```
Press CTRL+C to abort.
```

```
Input the bits of the modulus[default = 1024]:
```

```
Generating Keys...
```

```

+++++++
+++++++
+++++
+++++++
Generate a DSA key pair.
[Switch] public-key local create dsa
The range of public key size is (512 ~ 2048).
NOTES: If the key modulus is greater than 512,
It will take a few minutes.
Press CTRL+C to abort.
Input the bits of the modulus[default = 1024]:
Generating Keys...
+++++++
+++++++
Enable the SSH server.
[Switch] ssh server enable
Configure an IP address for VLAN-interface 1. This address will serve as the destination of the
SSH connection.
[Switch] interface vlan-interface 1
[Switch-Vlan-interface1] ip address 192.168.1.40 255.255.255.0
[Switch-Vlan-interface1] quit
Set the authentication mode for the user interfaces to AAA.
[Switch] user-interface vty 0 15
[Switch-ui-vty0-15] authentication-mode scheme
Enable the user interfaces to support SSH.
[Switch-ui-vty0-15] protocol inbound ssh
[Switch-ui-vty0-15] quit
Create local user client001, and set the user command privilege level to 3
[Switch] local-user client001
[Switch-luser-client001] password simple aabbcc
[Switch-luser-client001] service-type ssh
[Switch-luser-client001] authorization-attribute level 3
[Switch-luser-client001] quit
Specify the service type for user client001 as stelnet, and the authentication method as password.
This step is optional.
[Switch] ssh user client001 service-type stelnet authentication-type password

```

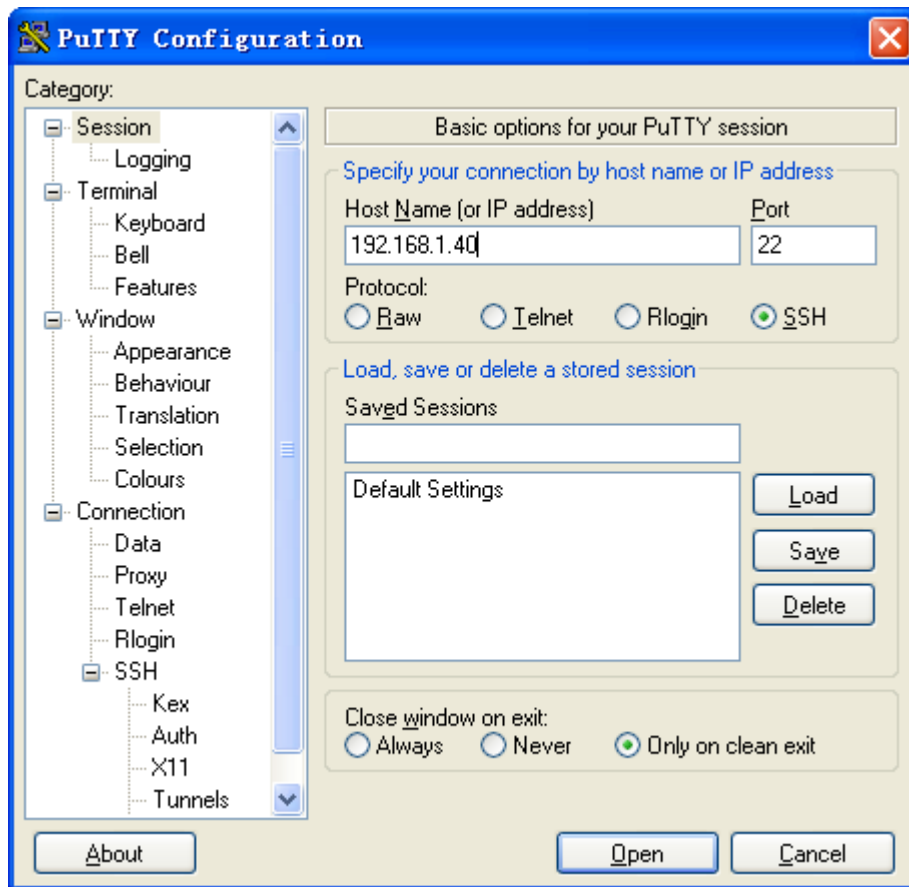
**2.** Establish a connection between the SSH client and the SSH server:

The switch supports a variety of SSH client software, such as PuTTY, and OpenSSH. The following example uses PuTTY Version 0.58.

# Establish a connection to the SSH server.

Launch PuTTY.exe to enter the following interface. In the **Host Name (or IP address)** text box, enter the IP address of the server (192.168.1.40).

Figure 102 Specifying the host name (or IP address)



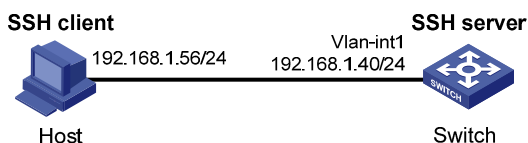
Click **Open** to connect to the server. If the connection is normal, you will be prompted to enter the username and password. After entering the username (**client001**) and password (**aabbcc**), you can enter the configuration interface of the server.

## When the switch acts as a server for publickey authentication

### Network requirements

As shown in Figure 103, a host (the SSH client) and a switch (the SSH server) are directly connected. Configure an SSH user on the switch so that the host can securely log in to the switch after passing publickey authentication. Use the RSA public key algorithm.

Figure 103 Network diagram



### Configuration procedure

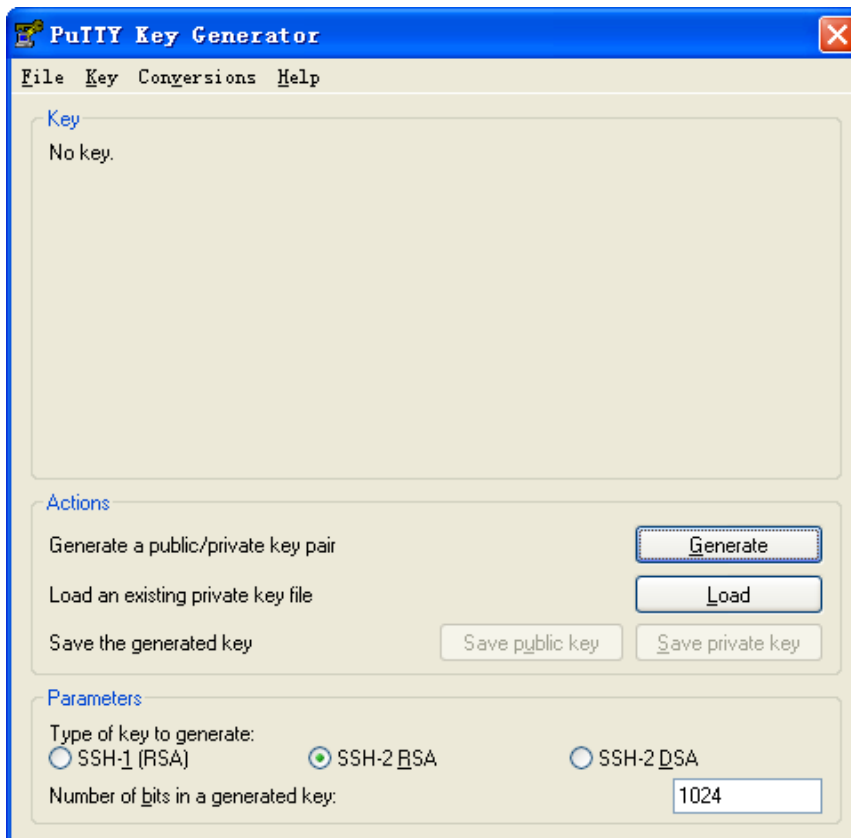


### IMPORTANT:

During SSH server configuration, the client public key is required. Use the client software to generate RSA key pairs on the client before configuring the SSH server.

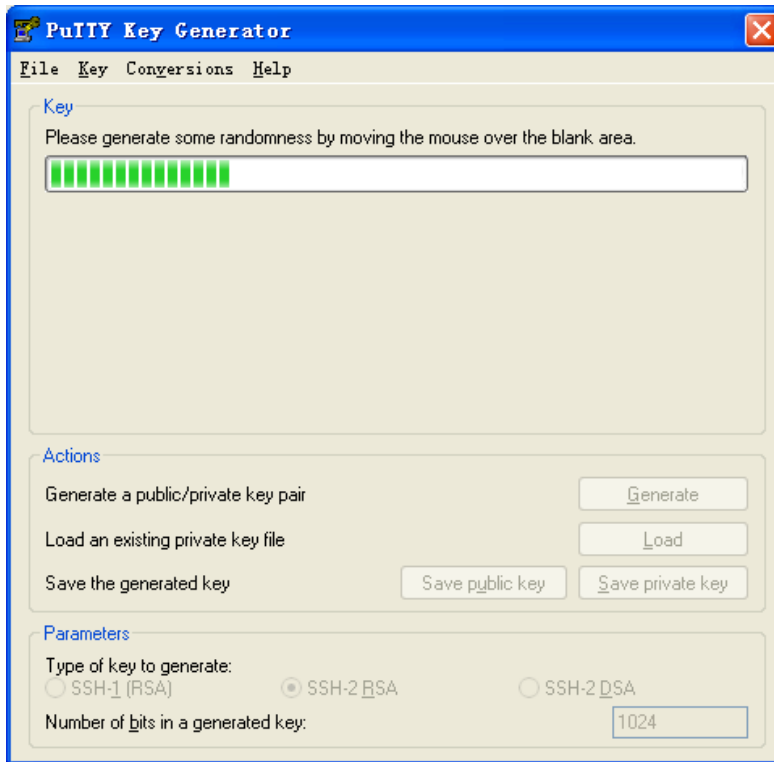
1. Configure the SSH client:
  - # Generate the RSA key pairs.
  - Run PuTTYGen.exe, select **SSH-2 RSA** and click **Generate**.

Figure 104 Generating the key pair on the client



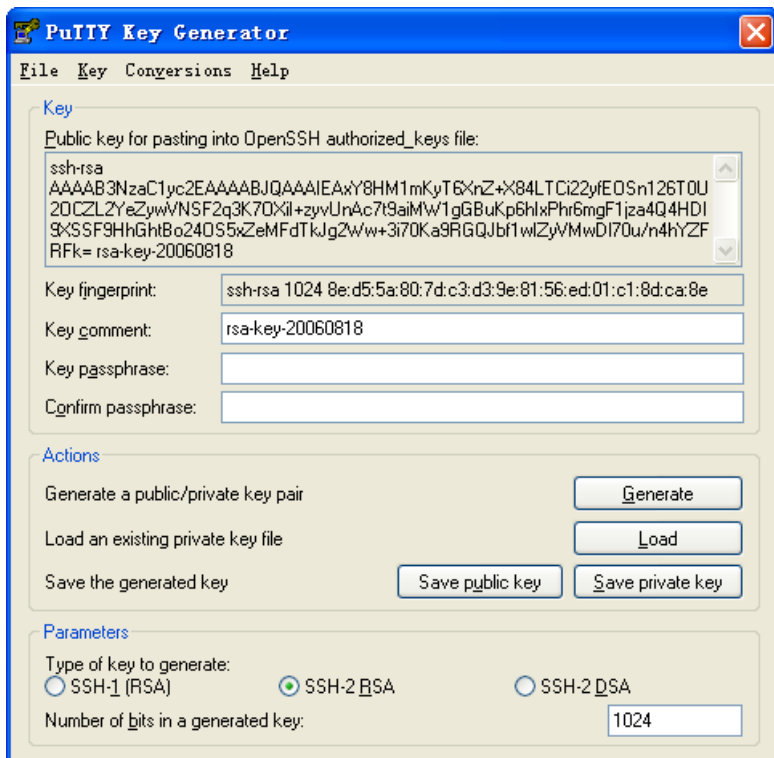
When the generator is generating the key pair, you must move the mouse continuously and keep the mouse off the green progress bar shown in Figure 105. Otherwise, the progress bar stops moving and the key pair generating process will be stopped.

Figure 105 Generating process



After the key pair is generated, click **Save public key** and specify the file name as **key.pub** to save the public key.

Figure 106 Saving the key pair on the client



Likewise, to save the private key, click **Save private key**. A warning window pops up to prompt you whether to save the private key without any protection. Click **Yes** and enter the name of the file for saving the key (**private.ppk** in this case).

Then, transmit the public key file to the server through FTP or TFTP.

## 2. Configure the SSH server:

# Generate the RSA key pairs.

```
<Switch> system-view
[Switch] public-key local create rsa
The range of public key size is (512 ~ 2048).
NOTES: If the key modulus is greater than 512,
It will take a few minutes.
Press CTRL+C to abort.
Input the bits of the modulus[default = 1024]:
Generating Keys...
+++++++
+++++
+++++
+++++
```

# Generate a DSA key pair.

```
[Switch] public-key local create dsa
The range of public key size is (512 ~ 2048).
NOTES: If the key modulus is greater than 512,
It will take a few minutes.
Press CTRL+C to abort.
Input the bits of the modulus[default = 1024]:
Generating Keys...
+++++
```

# Enable the SSH server.

```
[Switch] ssh server enable
```

# Configure an IP address for VLAN-interface 1. This address will serve as the destination of the SSH connection.

```
[Switch] interface vlan-interface 1
[Switch-Vlan-interface1] ip address 192.168.1.40 255.255.255.0
[Switch-Vlan-interface1] quit
```

# Set the authentication mode for the user interfaces to AAA.

```
[Switch] user-interface vty 0 15
[Switch-ui-vty0-15] authentication-mode scheme
```

# Enable the user interfaces to support SSH.

```
[Switch-ui-vty0-15] protocol inbound ssh
```

# Set the user command privilege level to 3.

```
[Switch-ui-vty0-15] user privilege level 3
[Switch-ui-vty0-15] quit
```

# Import the client's public key from file **key.pub** and name it **Switch001**.

```
[Switch] public-key peer Switch001 import sshkey key.pub
```

# Specify the authentication method for user **client002** as **publickey**, and assign the public key **Switch001** to the user.

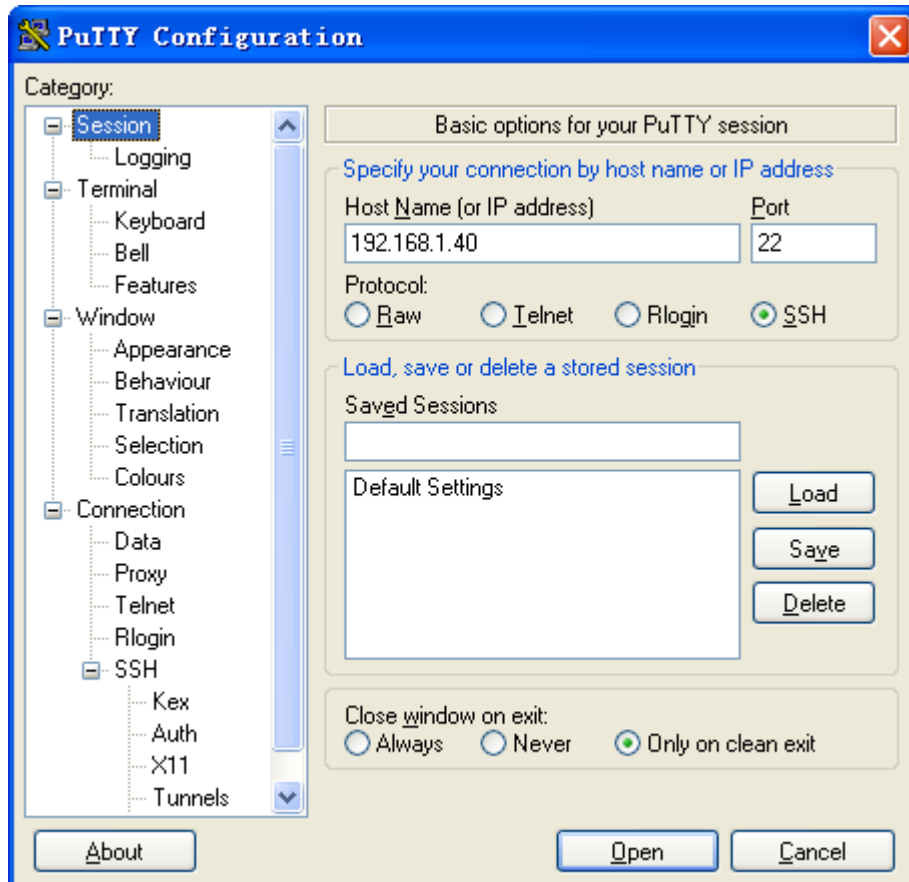
```
[Switch] ssh user client002 service-type stelnet authentication-type publickey assign publickey Switch001
```

3. Establish a connection between the SSH client and the SSH server:

# Specify the private key file and establish a connection to the SSH server

Launch PuTTY.exe to enter the following interface. In the **Host Name (or IP address)** text box, enter the IP address of the server (192.168.1.40).

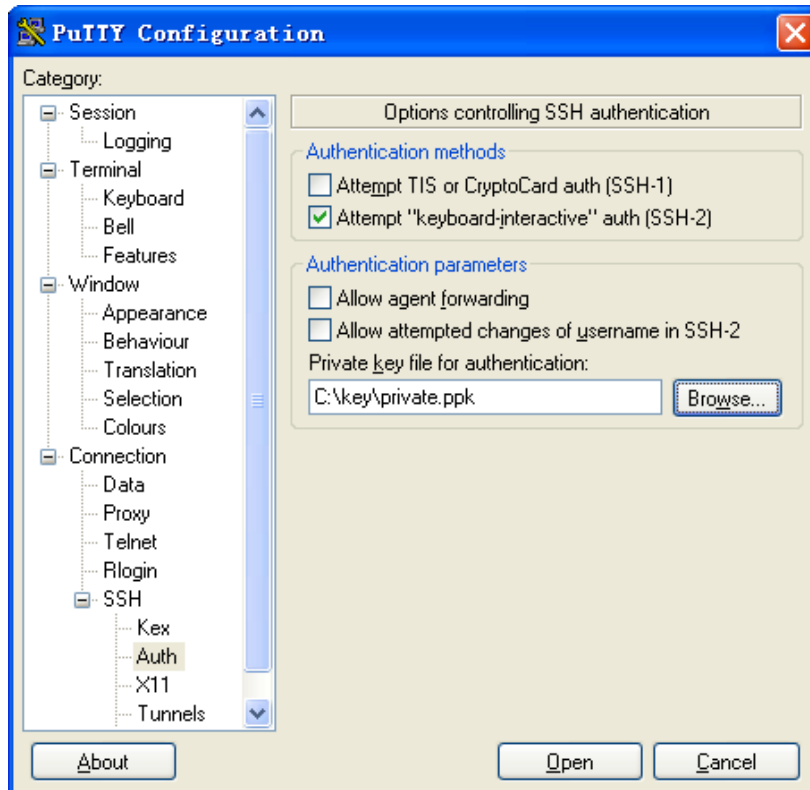
Figure 107 Specifying the host name (or IP address)



Select **Connection > SSH > Auth** from the navigation tree. The following window appears. Click **Browse...** to bring up the file selection window, navigate to the private key file (**private.ppk**) and click **OK**.



Figure 108 Specifying the private key file



Click **Open** to connect to the server. If the connection is normal, you will be prompted to enter the username. After entering the username (**client002**), you can enter the configuration interface of the server.

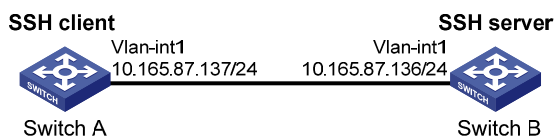
## SSH client configuration examples

### When switch acts as client for password authentication

#### Network requirements

As shown in Figure 109, Switch A (the SSH client) must pass password authentication to log in to Switch B (the SSH server) through the SSH protocol. Configure the username **client001** and the password **aabbcc** for the SSH client on Switch B.

Figure 109 Network diagram



#### Configuration procedure

1. Configure the SSH server:  
# Generate the RSA key pairs.  
<SwitchB> system-view

```

[SwitchB] public-key local create rsa
The range of public key size is (512 ~ 2048).
NOTES: If the key modulus is greater than 512,
It will take a few minutes.
Press CTRL+C to abort.
Input the bits of the modulus[default = 1024]:
Generating Keys...
+++++++
+++++
+++++
+++++
+++++
Generate a DSA key pair.
[SwitchB] public-key local create dsa
The range of public key size is (512 ~ 2048).
NOTES: If the key modulus is greater than 512,
It will take a few minutes.
Press CTRL+C to abort.
Input the bits of the modulus[default = 1024]:
Generating Keys...
+++++
+++++
Enable the SSH server.
[SwitchB] ssh server enable
Configure an IP address for VLAN-interface 1, which the SSH client will use as the destination for
SSH connection.
[SwitchB] interface vlan-interface 1
[SwitchB-Vlan-interface1] ip address 10.165.87.136 255.255.255.0
[SwitchB-Vlan-interface1] quit
Set the authentication mode for the user interfaces to AAA.
[SwitchB] user-interface vty 0 15
[SwitchB-ui-vty0-15] authentication-mode scheme
Enable the user interfaces to support SSH.
[SwitchB-ui-vty0-15] protocol inbound ssh
[SwitchB-ui-vty0-15] quit
Create local user client001.
[SwitchB] local-user client001
[SwitchB-luser-client001] password simple aabbcc
[SwitchB-luser-client001] service-type ssh
[SwitchB-luser-client001] authorization-attribute level 3
[SwitchB-luser-client001] quit
Specify the service type for user client001 as stelnet, and the authentication method as password.
This step is optional.
[SwitchB] ssh user client001 service-type stelnet authentication-type password

```

**2.** Establish a connection between the SSH client and the SSH server:

```

Configure an IP address for VLAN-interface 1.
<SwitchA> system-view
[SwitchA] interface vlan-interface 1

```

```
[SwitchA-Vlan-interface1] ip address 10.165.87.137 255.255.255.0
[SwitchA-Vlan-interface1] quit
[SwitchA] quit
```

- If the client supports first-time authentication, you can directly establish a connection from the client to the server.

# Establish an SSH connection to server 10.165.87.136.

```
<SwitchA> ssh2 10.165.87.136
Username: client001
Trying 10.165.87.136 ...
Press CTRL+K to abort
Connected to 10.165.87.136 ...
```

```
The Server is not authenticated. Continue? [Y/N]:y
Do you want to save the server public key? [Y/N]:n
Enter password:
```

After you enter the correct password, you can log in to Switch B successfully.

- If the client does not support first-time authentication, perform the following configurations.

# Disable first-time authentication.

```
[SwitchA] undo ssh client first-time
```

# Configure the host public key of the SSH server. You can get the server host public key by using the **display public-key local dsa public** command on the server.

```
[SwitchA] public-key peer key1
[SwitchA-pkey-public-key] public-key-code begin
[SwitchA-pkey-key-code]308201B73082012C06072A8648CE3804013082011F0281810
0D757262C4584C44C211F18BD96E5F0
[SwitchA-pkey-key-code]61C4F0A423F7FE6B6B85B34CEF72CE14A0D3A5222FE08CECE
65BE6C265854889DC1EDBD13EC8B274
[SwitchA-pkey-key-code]DA9F75BA26CCB987723602787E922BA84421F22C3C89CB9B0
6FD60FE01941DDD77FE6B12893DA76E
[SwitchA-pkey-key-code]EBC1D128D97F0678D7722B5341C8506F358214B16A2FAC4B3
68950387811C7DA33021500C773218C
[SwitchA-pkey-key-code]737EC8EE993B4F2DED30F48EDACE915F0281810082269009E
14EC474BAF2932E69D3B1F18517AD95
[SwitchA-pkey-key-code]94184CCDFCEAE96EC4D5EF93133E84B47093C52B20CD35D02
492B3959EC6499625BC4FA5082E22C5
[SwitchA-pkey-key-code]B374E16DD00132CE71B020217091AC717B612391C76C1FB2E
88317C1BD8171D41ECB83E210C03CC9
[SwitchA-pkey-key-code]B32E810561C21621C73D6DAAC028F4B1585DA7F42519718CC
9B09EEF0381840002818000AF995917
[SwitchA-pkey-key-code]E1E570A3F6B1C2411948B3B4FFA256699B3BF871221CC9C5D
F257523777D033BEE77FC378145F2AD
[SwitchA-pkey-key-code]D716D7DB9FCABB4ADBF6FB4FDB0CA25C761B308EF53009F71
01F7C62621216D5A572C379A32AC290
[SwitchA-pkey-key-code]E55B394A217DA38B65B77F0185C8DB8095522D1EF044B465E
8716261214A5A3B493E866991113B2D
[SwitchA-pkey-key-code]485348
[SwitchA-pkey-key-code] public-key-code end
```

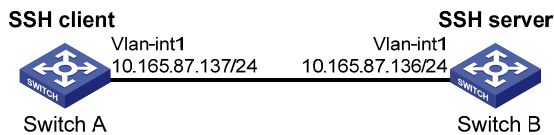
```
[SwitchA-pkey-public-key] peer-public-key end
Specify the host public key for the SSH server (10.165.87.136) as key1.
[SwitchA] ssh client authentication server 10.165.87.136 assign publickey key1
[SwitchA] quit
Establish an SSH connection to server 10.165.87.136.
<SwitchA> ssh2 10.165.87.136
Username: client001
Trying 10.165.87.136
Press CTRL+K to abort
Connected to 10.165.87.136...
Enter password:
After you enter the correct password, you can log in to Switch B successfully.
```

## When switch acts as client for publickey authentication

### Network requirements

As shown in [Figure 110](#), Switch A (the SSH client) must pass publickey authentication to log in to Switch B (the SSH server) through the SSH protocol. Use the DSA public key algorithm.

**Figure 110 Network diagram**



### Configuration procedure



#### IMPORTANT:

During SSH server configuration, the client public key is required. Use the client software to generate a DSA key pair on the client before configuring the SSH server.

#### 1. Configure the SSH client:

# Create VLAN-interface 1 and assign an IP address to it.

```
<SwitchA> system-view
[SwitchA] interface vlan-interface 1
[SwitchA-Vlan-interface1] ip address 10.165.87.137 255.255.255.0
[SwitchA-Vlan-interface1] quit
```

# Generate a DSA key pair.

```
[SwitchA] public-key local create dsa
```

The range of public key size is (512 ~ 2048).

NOTES: If the key modulus is greater than 512,

It will take a few minutes.

Press CTRL+C to abort.

Input the bits of the modulus[default = 1024]:

Generating Keys...

```
+++++
+++++
```

# Export the DSA public key to file **key.pub**.

```
[SwitchA] public-key local export dsa ssh2 key.pub
[SwitchA] quit
```

Then, transmit the public key file to the server through FTP or TFTP.

## 2. Configure the SSH server:

# Generate the RSA key pairs.

```
<SwitchB> system-view
[SwitchB] public-key local create rsa
The range of public key size is (512 ~ 2048).
NOTES: If the key modulus is greater than 512,
It will take a few minutes.
Press CTRL+C to abort.
Input the bits of the modulus[default = 1024]:
Generating Keys...
```

```
++++++
```

```
+++++
```

```
+++++
```

```
+++++
```

# Generate a DSA key pair.

```
[SwitchB] public-key local create dsa
The range of public key size is (512 ~ 2048).
NOTES: If the key modulus is greater than 512,
It will take a few minutes.
Press CTRL+C to abort.
Input the bits of the modulus[default = 1024]:
Generating Keys...
```

```
+++++
```

# Enable the SSH server.

```
[SwitchB] ssh server enable
```

# Configure an IP address for VLAN-interface 1, which the SSH client will use as the destination for SSH connection.

```
[SwitchB] interface vlan-interface 1
[SwitchB-Vlan-interface1] ip address 10.165.87.136 255.255.255.0
[SwitchB-Vlan-interface1] quit
```

# Set the authentication mode for the user interfaces to AAA.

```
[SwitchB] user-interface vty 0 15
[SwitchB-ui-vty0-15] authentication-mode scheme
```

# Enable the user interfaces to support SSH.

```
[SwitchB-ui-vty0-15] protocol inbound ssh
```

# Set the user command privilege level to 3.

```
[SwitchB-ui-vty0-15] user privilege level 3
[SwitchB-ui-vty0-15] quit
```

# Import the peer public key from the file **key.pub**.

```
[SwitchB] public-key peer Switch001 import sshkey key.pub
```

# Specify the authentication method for user **client002** as **publickey**, and assign the public key **Switch001** to the user.

```
[SwitchB] ssh user client002 service-type stelnet authentication-type publickey
assign publickey Switch001
```

**3.** Establish a connection between the SSH client and the SSH server:

# Establish an SSH connection to the server (10.165.87.136).

```
<SwitchA> ssh2 10.165.87.136
Username: client002
Trying 10.165.87.136 ...
Press CTRL+K to abort
Connected to 10.165.87.136 ...
```

```
The Server is not authenticated. Continue? [Y/N]:y
```

```
Do you want to save the server public key? [Y/N]:n
```

Later, you will find that you have logged in to Switch B successfully.

# Configuring SFTP

## Overview

The Secure File Transfer Protocol (SFTP) is a new feature in SSH2.0.

SFTP uses the SSH connection to provide secure data transfer. The switch can serve as the SFTP server, allowing a remote user to log in to the SFTP server for secure file management and transfer. The switch can also serve as an SFTP client, enabling a user to log in from the switch to a remote device for secure file transfer.

With SSH connection across VPNs, you can configure the switch as an SFTP client to establish connections with SFTP servers in different MPLS VPNs. For more information about this function, see "[Configuring SSH2.0](#)."

## Configuring the switch as an SFTP server

Before you configure this task, complete the following tasks:

- Configure the SSH server.
- Use the **ssh user service-type** command to set the service type of SSH users to **sftp** or **all**.

For more information about the configuration procedures, see "[Configuring SSH2.0](#)."

## Enabling the SFTP server

This configuration task will enable the SFTP service so that a client can log in to the SFTP server through SFTP.

When the switch functions as the SFTP server, only one client can access the SFTP server at a time. If the SFTP client uses WinSCP, a file on the server cannot be modified directly. It can only be downloaded to a local place, modified, and then uploaded to the server.

To enable the SFTP server:

| Step                       | Command                   | Remarks              |
|----------------------------|---------------------------|----------------------|
| 1. Enter system view.      | <b>system-view</b>        | N/A                  |
| 2. Enable the SFTP server. | <b>sftp server enable</b> | Disabled by default. |

## Configuring the SFTP connection idle timeout period

Once the idle period of an SFTP connection exceeds the specified threshold, the system automatically tears the connection down.

To configure the SFTP connection idle timeout period:

| Step                  | Command            | Remarks |
|-----------------------|--------------------|---------|
| 1. Enter system view. | <b>system-view</b> | N/A     |

| Step                                                  | Command                                                  | Remarks                             |
|-------------------------------------------------------|----------------------------------------------------------|-------------------------------------|
| 2. Configure the SFTP connection idle timeout period. | <b>sftp server idle-timeout</b><br><i>time-out-value</i> | Optional.<br>10 minutes by default. |

## Configuring the switch as an SFTP client

### Specifying a source IP address or interface for the SFTP client

You can configure a client to use only a specified source IP address or interface to access the SFTP server, enhancing the service manageability.

To specify a source IP address or interface for the SFTP client:

| Step                                                             | Command                                                                                                                                                                                                                                                                                                                                                                                     | Remarks                                                                                                                                                           |
|------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                                            | <b>system-view</b>                                                                                                                                                                                                                                                                                                                                                                          | N/A                                                                                                                                                               |
| 2. Specify a source IP address or interface for the SFTP client. | <ul style="list-style-type: none"> <li>Specify a source IPv4 address or interface for the SFTP client:<br/><b>sftp client source { ip ip-address   interface interface-type interface-number }</b></li> <li>Specify a source IPv6 address or interface for the SFTP client:<br/><b>sftp client ipv6 source { ipv6 ipv6-address   interface interface-type interface-number }</b></li> </ul> | <p>Use either command.</p> <p>By default, an SFTP client uses the IP address of the interface specified by the route of the switch to access the SFTP server.</p> |

### Establishing a connection to the SFTP server

This configuration task will enable the SFTP client to establish a connection to the remote SFTP server and enter SFTP client view.

To enable the SFTP client:



| Task                                                                         | Command                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | Remarks                                                                                                                                      |
|------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------|
| Establish a connection to the remote SFTP server and enter SFTP client view. | <ul style="list-style-type: none"> <li>Establish a connection to the remote IPv4 SFTP server and enter SFTP client view:<br/> <b>sftp server</b> [ <i>port-number</i> ] [ <b>vpn-instance</b> <i>vpn-instance-name</i> ] [ <b>identity-key</b> { <i>dsa</i>   <i>rsa</i> }   <b>prefer-ctos-cipher</b> { <i>3des</i>   <i>aes128</i>   <i>des</i> }   <b>prefer-ctos-hmac</b> { <i>md5</i>   <i>md5-96</i>   <i>sha1</i>   <i>sha1-96</i> }   <b>prefer-kex</b> { <i>dh-group-exchange</i>   <i>dh-group1</i>   <i>dh-group14</i> }   <b>prefer-stoc-cipher</b> { <i>3des</i>   <i>aes128</i>   <i>des</i> }   <b>prefer-stoc-hmac</b> { <i>md5</i>   <i>md5-96</i>   <i>sha1</i>   <i>sha1-96</i> } ] *</li> <li>Establish a connection to the remote IPv6 SFTP server and enter SFTP client view:<br/> <b>sftp ipv6 server</b> [ <i>port-number</i> ] [ <b>vpn-instance</b> <i>vpn-instance-name</i> ] [ <b>identity-key</b> { <i>dsa</i>   <i>rsa</i> }   <b>prefer-ctos-cipher</b> { <i>3des</i>   <i>aes128</i>   <i>des</i> }   <b>prefer-ctos-hmac</b> { <i>md5</i>   <i>md5-96</i>   <i>sha1</i>   <i>sha1-96</i> }   <b>prefer-kex</b> { <i>dh-group-exchange</i>   <i>dh-group1</i>   <i>dh-group14</i> }   <b>prefer-stoc-cipher</b> { <i>3des</i>   <i>aes128</i>   <i>des</i> }   <b>prefer-stoc-hmac</b> { <i>md5</i>   <i>md5-96</i>   <i>sha1</i>   <i>sha1-96</i> } ] *</li> </ul> | <p>Use either command in user view.</p> <p>Only the HP 5500 EI switches support the <b>vpn-instance</b> <i>vpn-instance-name</i> option.</p> |

## Working with SFTP directories

SFTP directory operations include:

- Changing or displaying the current working directory
- Displaying files under a directory or the directory information
- Changing the name of a directory on the server
- Creating or deleting a directory

To work with the SFTP directories:

| Step                                                                | Command                                                                                                                                                                               | Remarks                                                                 |
|---------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------|
| 1. Enter SFTP client view.                                          | For more information, see " <a href="#">Establishing a connection to the SFTP server.</a> "                                                                                           | Execute the command in user view.                                       |
| 2. Change the working directory of the remote SFTP server.          | <b>cd</b> [ <i>remote-path</i> ]                                                                                                                                                      | Optional.                                                               |
| 3. Return to the upper-level directory.                             | <b>cdup</b>                                                                                                                                                                           | Optional.                                                               |
| 4. Display the current working directory of the remote SFTP server. | <b>pwd</b>                                                                                                                                                                            | Optional.                                                               |
| 5. Display files under a directory.                                 | <ul style="list-style-type: none"> <li>• <b>dir</b> [ <b>-a</b>   <b>-l</b> ] [ <i>remote-path</i> ]</li> <li>• <b>ls</b> [ <b>-a</b>   <b>-l</b> ] [ <i>remote-path</i> ]</li> </ul> | Optional.<br>The <b>dir</b> command functions as the <b>ls</b> command. |
| 6. Change the name of a directory on the SFTP server.               | <b>rename</b> <i>oldname newname</i>                                                                                                                                                  | Optional.                                                               |

| Step                                                    | Command                                 | Remarks   |
|---------------------------------------------------------|-----------------------------------------|-----------|
| 7. Create a new directory on the remote SFTP server.    | <b>mkdir</b> <i>remote-path</i>         | Optional. |
| 8. Delete one or more directories from the SFTP server. | <b>rmdir</b> <i>remote-path</i> &<1-10> | Optional. |

## Working with SFTP files

SFTP file operations include:

- Changing the name of a file
- Downloading a file
- Uploading a file
- Displaying a list of the files
- Deleting a file

To work with SFTP files:

| Step                                                           | Command                                                                                                                                                                               | Remarks                                                                        |
|----------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|
| 1. Enter SFTP client view.                                     | For more information, see " <a href="#">Establishing a connection to the SFTP server.</a> "                                                                                           | Execute the command in user view.                                              |
| 2. Change the name of a file on the SFTP server.               | <b>rename</b> <i>old-name new-name</i>                                                                                                                                                | Optional.                                                                      |
| 3. Download a file from the remote server and save it locally. | <b>get</b> <i>remote-file</i> [ <i>local-file</i> ]                                                                                                                                   | Optional.                                                                      |
| 4. Upload a local file to the remote SFTP server.              | <b>put</b> <i>local-file</i> [ <i>remote-file</i> ]                                                                                                                                   | Optional.                                                                      |
| 5. Display the files under a directory.                        | <ul style="list-style-type: none"> <li>• <b>dir</b> [ <b>-a</b>   <b>-l</b> ] [ <i>remote-path</i> ]</li> <li>• <b>ls</b> [ <b>-a</b>   <b>-l</b> ] [ <i>remote-path</i> ]</li> </ul> | Optional.<br>The <b>dir</b> command functions as the <b>ls</b> command.        |
| 6. Delete one or more directories from the SFTP server.        | <ul style="list-style-type: none"> <li>• <b>delete</b> <i>remote-file</i>&amp;&lt;1-10&gt;</li> <li>• <b>remove</b> <i>remote-file</i>&amp;&lt;1-10&gt;</li> </ul>                    | Optional.<br>The <b>delete</b> command functions as the <b>remove</b> command. |

## Displaying help information

This configuration task will display a list of all commands or the help information of an SFTP client command, such as the command format and parameters.

To display a list of all commands or the help information of an SFTP client command:

| Step                       | Command                                                                                     | Remarks                           |
|----------------------------|---------------------------------------------------------------------------------------------|-----------------------------------|
| 1. Enter SFTP client view. | For more information, see " <a href="#">Establishing a connection to the SFTP server.</a> " | Execute the command in user view. |

| Step | Command                                                                           | Remarks                                         |
|------|-----------------------------------------------------------------------------------|-------------------------------------------------|
| 2.   | Display a list of all commands or the help information of an SFTP client command. | <code>help [ all   command-name ]</code><br>N/A |

## Terminating the connection to the remote SFTP server

| Step | Command                                                                     | Remarks                                                                                                                                                                                             |
|------|-----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1.   | Enter SFTP client view.                                                     | For more information, see " <a href="#">Establishing a connection to the SFTP server.</a> "<br>Execute the command in user view.                                                                    |
| 2.   | Terminate the connection to the remote SFTP server and return to user view. | <ul style="list-style-type: none"> <li><code>bye</code></li> <li><code>exit</code></li> <li><code>quit</code></li> </ul> Use any of the commands.<br>These three commands function in the same way. |

## Setting the DSCP value for packets sent by the SFTP client

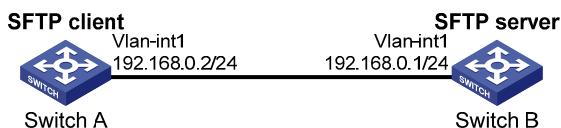
| Step | Command                                                 | Remarks                                                                                                                                                                                                                                                                                                                                                                                                                     |
|------|---------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1.   | Enter system view.                                      | <code>system-view</code><br>N/A                                                                                                                                                                                                                                                                                                                                                                                             |
| 2.   | Set the DSCP value for packets sent by the SFTP client. | <ul style="list-style-type: none"> <li>Set the DSCP value for packets sent by the IPv4 SFTP client:<br/><code>sftp client dscp dscp-value</code></li> <li>Set the DSCP value for packets sent by the IPv6 SFTP client:<br/><code>sftp client ipv6 dscp dscp-value</code></li> </ul> Optional.<br>By default, the DSCP value is 16 in packets sent by the IPv4 SFTP client and is 8 in packets sent by the IPv6 SFTP client. |

## SFTP client configuration example

### Network requirements

As shown in [Figure 111](#), an SSH connection is required between Switch A and Switch B. Switch A, an SFTP client, needs to log in to Switch B for file management and file transfer. Use publickey authentication and the RSA public key algorithm.

**Figure 111 Network diagram**



### Configuration procedure



---

**IMPORTANT:**

During SFTP server configuration, the client public key is required. Use the client software to generate RSA key pairs on the client before configuring the SFTP server.

---

**1. Configure the SFTP client:**

# Create VLAN-interface 1 and assign an IP address to it.

```
<SwitchA> system-view
[SwitchA] interface vlan-interface 1
[SwitchA-Vlan-interface1] ip address 192.168.0.2 255.255.255.0
[SwitchA-Vlan-interface1] quit
```

# Generate the RSA key pairs.

```
[SwitchA] public-key local create rsa
The range of public key size is (512 ~ 2048).
NOTES: If the key modulus is greater than 512,
It will take a few minutes.
Press CTRL+C to abort.
```

```
Input the bits of the modulus[default = 1024]:
```

```
Generating Keys...
```

```
++++++
```

```
+++++
```

```
+++++
```

```
+++++
```

# Export the host public key to file **pubkey**.

```
[SwitchA] public-key local export rsa ssh2 pubkey
[SwitchA] quit
```

Then, transmit the public key file to the server through FTP or TFTP.

**2. Configure the SFTP server:**

# Generate the RSA key pairs.

```
<SwitchB> system-view
[SwitchB] public-key local create rsa
The range of public key size is (512 ~ 2048).
NOTES: If the key modulus is greater than 512,
It will take a few minutes.
Press CTRL+C to abort.
```

```
Input the bits of the modulus[default = 1024]:
```

```
Generating Keys...
```

```
++++++
```

```
+++++
```

```
+++++
```

```
+++++
```

# Generate a DSA key pair.

```
[SwitchB] public-key local create dsa
The range of public key size is (512 ~ 2048).
NOTES: If the key modulus is greater than 512,
It will take a few minutes.
Press CTRL+C to abort.
```

```
Input the bits of the modulus[default = 1024]:
```

Generating Keys...

```
+++++
```

# Enable the SSH server.

```
[SwitchB] ssh server enable
```

# Enable the SFTP server.

```
[SwitchB] sftp server enable
```

# Configure an IP address for VLAN-interface 1, which the SSH client uses as the destination for SSH connection.

```
[SwitchB] interface vlan-interface 1
```

```
[SwitchB-Vlan-interface1] ip address 192.168.0.1 255.255.255.0
```

```
[SwitchB-Vlan-interface1] quit
```

# Set the authentication mode on the user interfaces to AAA.

```
[SwitchB] user-interface vty 0 15
```

```
[SwitchB-ui-vty0-15] authentication-mode scheme
```

# Set the protocol that a remote user uses to log in as **SSH**.

```
[SwitchB-ui-vty0-15] protocol inbound ssh
```

```
[SwitchB-ui-vty0-15] quit
```

# Import the peer public key from the file **pubkey**.

```
[SwitchB] public-key peer Switch001 import sshkey pubkey
```

# For user **client001**, set the service type as SFTP, authentication method as publickey, public key as **Switch001**, and working folder as **flash:/**

```
[SwitchB] ssh user client001 service-type sftp authentication-type publickey assign publickey Switch001 work-directory flash:/
```

### 3. Establish a connection between the SFTP client and the SFTP server:

# Establish a connection to the remote SFTP server and enter SFTP client view.

```
<SwitchA> sftp 192.168.0.1 identity-key rsa
```

```
Input Username: client001
```

```
Trying 192.168.0.1 ...
```

```
Press CTRL+K to abort
```

```
Connected to 192.168.0.1 ...
```

```
The Server is not authenticated. Continue? [Y/N]:y
```

```
Do you want to save the server public key? [Y/N]:n
```

```
sftp-client>
```

# Display files under the current directory of the server, delete the file named **z**, and check if the file has been deleted successfully.

```
sftp-client> dir
```

```
-rwxrwxrwx 1 noone nogroup 1759 Aug 23 06:52 config.cfg
```

```
-rwxrwxrwx 1 noone nogroup 225 Aug 24 08:01 pubkey2
```

```
-rwxrwxrwx 1 noone nogroup 283 Aug 24 07:39 pubkey
```

```
drwxrwxrwx 1 noone nogroup 0 Sep 01 06:22 new
```

```
-rwxrwxrwx 1 noone nogroup 225 Sep 01 06:55 pub
```

```
-rwxrwxrwx 1 noone nogroup 0 Sep 01 08:00 z
```

```
sftp-client> delete z
```

```
The following File will be deleted:
```

```
/z
```

```
Are you sure to delete it? [Y/N]:y
```

```
This operation might take a long time.Please wait...
```

```
File successfully Removed
```

```
sftp-client> dir
```

```
-rwxrwxrwx 1 noone nogroup 1759 Aug 23 06:52 config.cfg
-rwxrwxrwx 1 noone nogroup 225 Aug 24 08:01 pubkey2
-rwxrwxrwx 1 noone nogroup 283 Aug 24 07:39 pubkey
drwxrwxrwx 1 noone nogroup 0 Sep 01 06:22 new
-rwxrwxrwx 1 noone nogroup 225 Sep 01 06:55 pub
```

```
Add a directory named new1 and check if it has been created successfully.
```

```
sftp-client> mkdir new1
```

```
New directory created
```

```
sftp-client> dir
```

```
-rwxrwxrwx 1 noone nogroup 1759 Aug 23 06:52 config.cfg
-rwxrwxrwx 1 noone nogroup 225 Aug 24 08:01 pubkey2
-rwxrwxrwx 1 noone nogroup 283 Aug 24 07:39 pubkey
drwxrwxrwx 1 noone nogroup 0 Sep 01 06:22 new
-rwxrwxrwx 1 noone nogroup 225 Sep 01 06:55 pub
drwxrwxrwx 1 noone nogroup 0 Sep 02 06:30 new1
```

```
Rename directory new1 to new2 and check if the directory has been renamed successfully.
```

```
sftp-client> rename new1 new2
```

```
File successfully renamed
```

```
sftp-client> dir
```

```
-rwxrwxrwx 1 noone nogroup 1759 Aug 23 06:52 config.cfg
-rwxrwxrwx 1 noone nogroup 225 Aug 24 08:01 pubkey2
-rwxrwxrwx 1 noone nogroup 283 Aug 24 07:39 pubkey
drwxrwxrwx 1 noone nogroup 0 Sep 01 06:22 new
-rwxrwxrwx 1 noone nogroup 225 Sep 01 06:55 pub
drwxrwxrwx 1 noone nogroup 0 Sep 02 06:33 new2
```

```
Download the pubkey2 file from the server and save it as local file public.
```

```
sftp-client> get pubkey2 public
```

```
Remote file:/pubkey2 ---> Local file: public
```

```
Downloading file successfully ended
```

```
Upload the local file pu to the server, save it as puk, and check if the file has been uploaded successfully.
```

```
sftp-client> put pu puk
```

```
Local file:pu ---> Remote file: /puk
```

```
Uploading file successfully ended
```

```
sftp-client> dir
```

```
-rwxrwxrwx 1 noone nogroup 1759 Aug 23 06:52 config.cfg
-rwxrwxrwx 1 noone nogroup 225 Aug 24 08:01 pubkey2
-rwxrwxrwx 1 noone nogroup 283 Aug 24 07:39 pubkey
drwxrwxrwx 1 noone nogroup 0 Sep 01 06:22 new
drwxrwxrwx 1 noone nogroup 0 Sep 02 06:33 new2
-rwxrwxrwx 1 noone nogroup 283 Sep 02 06:35 pub
-rwxrwxrwx 1 noone nogroup 283 Sep 02 06:36 puk
```

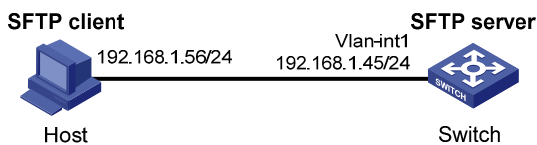
```
sftp-client>
Terminate the connection to the remote SFTP server.
sftp-client> quit
Bye
Connection closed.
<SwitchA>
```

## SFTP server configuration example

### Network requirements

As shown in [Figure 112](#), an SSH connection is required between the host and the switch. The host, an SFTP client, needs to log in to the switch for file management and file transfer. Use password authentication and configure the username **client002** and the password **aabbcc** for the client on the switch.

**Figure 112 Network diagram**



### Configuration procedure

1. Configure the SFTP server:

```
Generate the RSA key pairs.
```

```
<Switch> system-view
[Switch] public-key local create rsa
The range of public key size is (512 ~ 2048).
NOTES: If the key modulus is greater than 512,
It will take a few minutes.
Press CTRL+C to abort.
Input the bits of the modulus[default = 1024]:
Generating Keys...
+++++++
++++
++++
++++
++++
++++
Generate a DSA key pair.
[Switch] public-key local create dsa
The range of public key size is (512 ~ 2048).
NOTES: If the key modulus is greater than 512,
It will take a few minutes.
Press CTRL+C to abort.
Input the bits of the modulus[default = 1024]:
Generating Keys...
+++++
```

```
Enable the SSH server.
```

```

[Switch] ssh server enable
Enable the SFTP server.
[Switch] sftp server enable
Configure an IP address for VLAN-interface 1, which the client will use as the destination for SSH
connection.
[Switch] interface vlan-interface 1
[Switch-Vlan-interface1] ip address 192.168.1.45 255.255.255.0
[Switch-Vlan-interface1] quit
Set the authentication mode of the user interfaces to AAA.
[Switch] user-interface vty 0 15
[Switch-ui-vty0-15] authentication-mode scheme
Enable the user interfaces to support SSH.
[Switch-ui-vty0-15] protocol inbound ssh
[Switch-ui-vty0-15] quit
Configure a local user named client002 with the password being aabbcc and the service type
being SSH.
[Switch] local-user client002
[Switch-luser-client002] password simple aabbcc
[Switch-luser-client002] service-type ssh
[Switch-luser-client002] quit
Configure the user authentication method as password and service type as SFTP.
[Switch] ssh user client002 service-type sftp authentication-type password

```

**2.** Establish a connection between the SFTP client and the SFTP server:

The switch supports a variety of SFTP client software. The following example uses PSFTP of PuTTY Version 0.58.

---

**NOTE:**

PSFTP supports only password authentication.

---

# Establish a connection to the remote SFTP server.

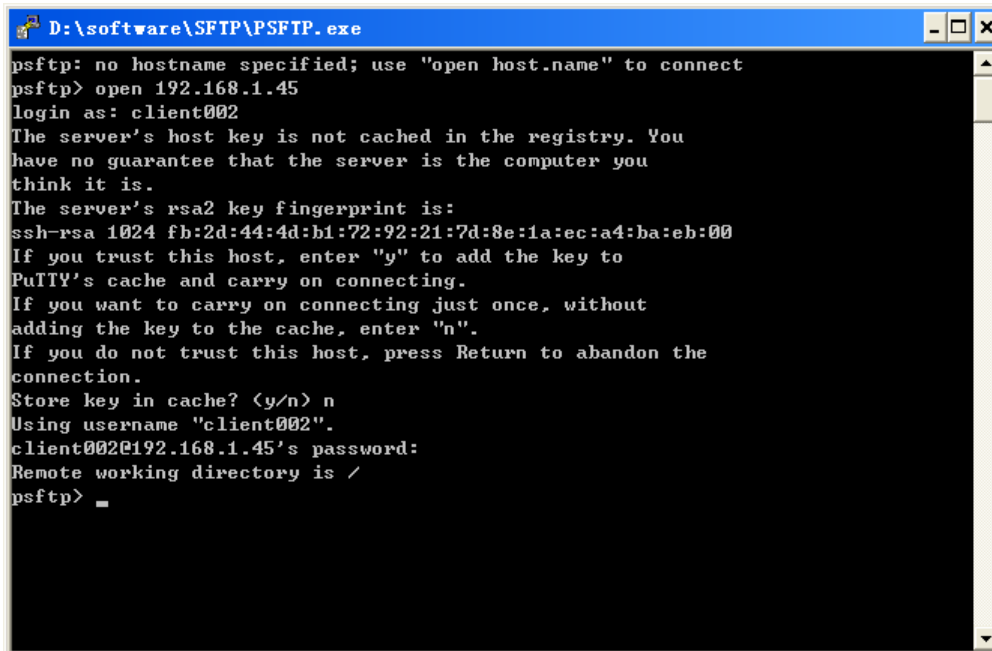
Run the psftp.exe to launch the client interface as shown in [Figure 113](#), and enter the following command:

```
open 192.168.1.45
```

Enter username **client002** and password **aabbcc** as prompted to log in to the SFTP server.



Figure 113 SFTP client interface



```
D:\software\SFTP\PSFTP.exe
psftp: no hostname specified; use "open host.name" to connect
psftp> open 192.168.1.45
login as: client002
The server's host key is not cached in the registry. You
have no guarantee that the server is the computer you
think it is.
The server's rsa2 key fingerprint is:
ssh-rsa 1024 fb:2d:44:4d:b1:72:92:21:7d:8e:1a:ec:a4:ba:eb:00
If you trust this host, enter "y" to add the key to
PuTTY's cache and carry on connecting.
If you want to carry on connecting just once, without
adding the key to the cache, enter "n".
If you do not trust this host, press Return to abandon the
connection.
Store key in cache? (y/n) n
Using username "client002".
client002@192.168.1.45's password:
Remote working directory is /
psftp> _
```

# Configuring SCP

## Overview

Secure copy (SCP) is based on SSH2.0 and offers a secure approach to copying files.

SCP uses SSH connections for copying files. The switch can act as the SCP server, allowing a user to log in to the switch for file upload and download. The switch can also act as an SCP client, enabling a user to log in from the switch to a remote server for secure file transfer.

### NOTE:

When the switch acts as an SCP server, only one of the FTP, SFTP or SCP user can access the switch.

## Configuring the switch as an SCP server

| Step                                                                                                                                          | Command                                                                                                                                                                                                                                                                                                                                      | Remarks                                                                                                                                                                                                                                      |
|-----------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                                                                                                                         | <b>system-view</b>                                                                                                                                                                                                                                                                                                                           | N/A                                                                                                                                                                                                                                          |
| 2. Configure the SSH server.                                                                                                                  | For more information, see the security guide for your switch.                                                                                                                                                                                                                                                                                | N/A                                                                                                                                                                                                                                          |
| 3. Create an SSH user for a SCP client, set the service type to <b>all</b> or <b>scp</b> , and specify the authentication method.             | <b>ssh user</b> <i>username</i> <b>service-type</b> { <b>all</b>   <b>scp</b> }<br><b>authentication-type</b> { <b>password</b>   { <b>any</b>   <b>password-publickey</b>   <b>publickey</b> } <b>assign</b><br><b>publickey</b> <i>keyname</i> <b>work-directory</b> <i>directory-name</i> }                                               | N/A                                                                                                                                                                                                                                          |
| 4. Create a user account and assign a working directory for the SSH user on the switch or a remote server if password authentication is used. | <ul style="list-style-type: none"><li>On the remote server (Details not shown.)</li><li>On the switch:<ul style="list-style-type: none"><li><b>a. local-user</b></li><li><b>b. password</b></li><li><b>c. service-type ssh</b></li><li><b>d. authorization-attribute</b><br/><b>work-directory</b> <i>directory-name</i></li></ul></li></ul> | <p>Skip this step if publickey authentication, whether with password authentication or not, is used.</p> <p>Make sure that the local user account has the name <i>username</i> as the username specified in the <b>ssh user</b> command.</p> |

When you set the working directory for the user, follow these guidelines:

- If only password authentication is used, the working directory specified in the **ssh user** command does not take effect. You must set the working directory on the remote server or in the local user account for the SSH user.
- If publickey authentication, whether with password authentication or not, is used, you must set the working directory in the **ssh user** command.

## Configuring the switch as the SCP client

To upload or download files to or from an SCP server:

| Step                                   | Command                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Remarks                  |
|----------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------|
| 1. Upload a file to an SCP server.     | <ul style="list-style-type: none"> <li>Upload a file to the IPv4 SCP server:<br/> <code>scp server [ port-number ] put source-file-path [ destination-file-path ] [ identity-key { dsa   rsa }   prefer-ctos-cipher { 3des   aes128   des }   prefer-ctos-hmac { md5   md5-96   sha1   sha1-96 }   prefer-kex { dh-group-exchange   dh-group1   dh-group14 }   prefer-stoc-cipher { 3des   aes128   des }   prefer-stoc-hmac { md5   md5-96   sha1   sha1-96 } ] *</code></li> <li>Upload a file to the IPv6 SCP server:<br/> <code>scp ipv6 server [ port-number ] put source-file-path [ destination-file-path ] [ identity-key { dsa   rsa }   prefer-ctos-cipher { 3des   aes128   des }   prefer-ctos-hmac { md5   md5-96   sha1   sha1-96 }   prefer-kex { dh-group-exchange   dh-group1   dh-group14 }   prefer-stoc-cipher { 3des   aes128   des }   prefer-stoc-hmac { md5   md5-96   sha1   sha1-96 } ] *</code></li> </ul>                       | Use one of the commands. |
| 2. Download a file from an SCP server. | <ul style="list-style-type: none"> <li>Download a file from the remote IPv4 SCP server:<br/> <code>scp server [ port-number ] get source-file-path [ destination-file-path ] [ identity-key { dsa   rsa }   prefer-ctos-cipher { 3des   aes128   des }   prefer-ctos-hmac { md5   md5-96   sha1   sha1-96 }   prefer-kex { dh-group-exchange   dh-group1   dh-group14 }   prefer-stoc-cipher { 3des   aes128   des }   prefer-stoc-hmac { md5   md5-96   sha1   sha1-96 } ] *</code></li> <li>Download a file from the remote IPv6 SCP server:<br/> <code>scp ipv6 server [ port-number ] get source-file-path [ destination-file-path ] [ identity-key { dsa   rsa }   prefer-ctos-cipher { 3des   aes128   des }   prefer-ctos-hmac { md5   md5-96   sha1   sha1-96 }   prefer-kex { dh-group-exchange   dh-group1   dh-group14 }   prefer-stoc-cipher { 3des   aes128   des }   prefer-stoc-hmac { md5   md5-96   sha1   sha1-96 } ] *</code></li> </ul> | Available in user view.  |



#### IMPORTANT:

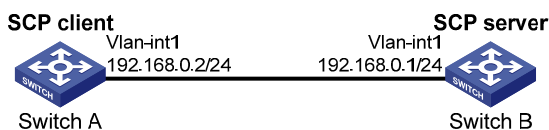
File transfer interruption during a downloading process can result in file fragments on the switch. You must manually delete them.

## SCP client configuration example

### Network requirements

As shown in [Figure 114](#), switch A acts as a client and download the file **remote.bin** from switch B. The user has the username **test** and uses the password authentication method.

**Figure 114 Network diagram**



### Configuration procedure

# Create VLAN-interface 1 and assign an IP address to it.

```

<SwitchA> system-view
[SwitchA] interface vlan-interface 1
[SwitchA-Vlan-interface1] ip address 192.168.0.2 255.255.255.0
[SwitchA-Vlan-interface1] quit

```

# Download the file **remote.bin** from the SCP server, save it locally and change the file name to **local.bin**.

```

<SwitchA> scp 192.168.0.1 get remote.bin local.bin
Username: test
Trying 192.168.0.1 ...
Press CTRL+K to abort
Connected to 192.168.0.1 ...

```

```

The Server is not authenticated. Continue? [Y/N]:y
Do you want to save the server public key? [Y/N]:n
Enter password:
18471 bytes transferred in 0.001 seconds.

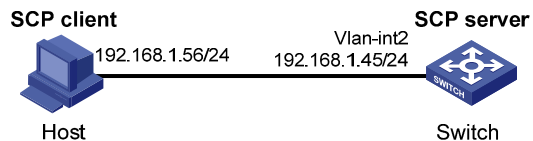
```

## SCP server configuration example

### Network requirements

As shown in [Figure 115](#), the switch acts as the SCP server, and the host acts as the SCP client. The host establishes an SSH connection to the switch. The user uses the username **test** and the password **aabbcc**. The username and password are saved on the switch for local authentication.

**Figure 115 Network diagram**



### Configuration procedure

# Generate the RSA key pairs.

```

<Switch> system-view
[Switch] public-key local create rsa
The range of public key size is (512 ~ 2048).
NOTES: If the key modulus is greater than 512,
It will take a few minutes.
Press CTRL+C to abort.
Input the bits of the modulus[default = 1024]:
Generating Keys...
+++++++
+++++
+++++
+++++
+++++++

```

# Generate the DSA key pair.

```

[Switch] public-key local create dsa
The range of public key size is (512 ~ 2048).
NOTES: If the key modulus is greater than 512,

```

```

It will take a few minutes.
Press CTRL+C to abort.
Input the bits of the modulus[default = 1024]:
Generating Keys...
+++++
+++++

Enable the SSH server function.
[Switch] ssh server enable

Configure an IP address for VLAN-interface 1, which the client will use as the destination for SSH
connection.
[Switch] interface vlan-interface 1
[Switch-Vlan-interface1] ip address 192.168.1.45 255.255.255.0
[Switch-Vlan-interface1] quit

Set the authentication mode of the user interfaces to AAA.
[Switch] user-interface vty 0 15
[Switch-ui-vty0-15] authentication-mode scheme

Enable the user interfaces to support all protocols including SSH.
[Switch-ui-vty0-15] protocol inbound all
[Switch-ui-vty0-15] quit

Create a local user named test.
[Switch] local-user test
[Switch-luser-test] password simple aabbcc
[Switch-luser-test] service-type ssh
[Switch-luser-test] quit

Configure the SSH user authentication method as password and service type as scp.
[Switch] ssh user test service-type scp authentication-type password

```

# Configuring SSL

## Overview

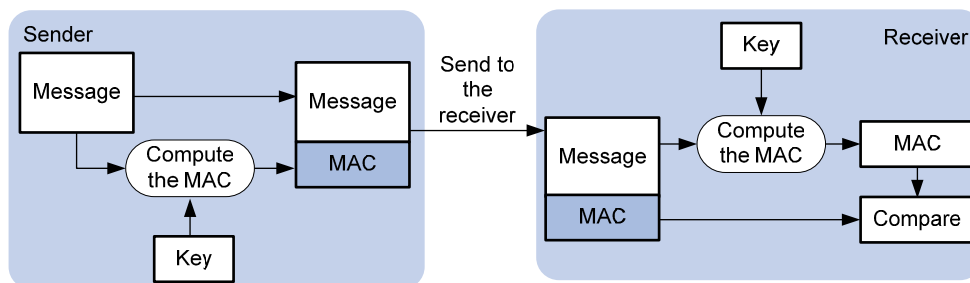
Secure Sockets Layer (SSL) is a security protocol that provides secure connection services for TCP-based application layer protocols such as Hypertext Transfer Protocol (HTTP). It is widely used in e-business and online banking to ensure secure data transmission over the Internet.

## SSL security mechanism

Secure connections provided by SSL have these features:

- **Confidentiality**—SSL uses a symmetric encryption algorithm to encrypt data and uses the asymmetric key algorithm of Rivest, Shamir, and Adelman (RSA) to encrypt the key to be used by the symmetric encryption algorithm.
- **Authentication**—SSL supports certificate-based identity authentication of the server and client by using the digital signatures. The SSL server and client obtain certificates from a certificate authority (CA) through the Public Key Infrastructure (PKI).
- **Reliability**—SSL uses the key-based message authentication code (MAC) to verify message integrity. A MAC algorithm transforms a message of any length to a fixed-length message. With the key, the sender uses the MAC algorithm to compute the MAC value of a message. Then, the sender suffixes the MAC value to the message and sends the result to the receiver. The receiver uses the same key and MAC algorithm to compute the MAC value of the received message, and compares the locally computed MAC value with that received. If the two values match, the receiver considers the message intact; otherwise, the receiver considers that the message has been tampered with in transit and discards the message.

Figure 116 Message integrity verification by a MAC algorithm



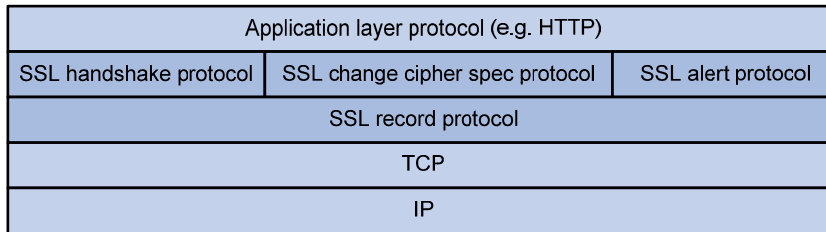
For more information about symmetric key algorithms, asymmetric key algorithm RSA and digital signature, see "[Managing public keys.](#)"

For more information about PKI, certificate, and CA, see "[Configuring PKI.](#)"

## SSL protocol stack

The SSL protocol consists of two layers of protocols: the SSL record protocol at the lower layer and the SSL handshake protocol, change cipher spec protocol, and alert protocol at the upper layer.

**Figure 117 SSL protocol stack**



- **SSL record protocol**—Fragments data to be transmitted, computes and adds MAC to the data, and encrypts the data before transmitting it to the peer end.
- **SSL handshake protocol**—Negotiates the cipher suite to be used for secure communication (including the symmetric encryption algorithm, key exchange algorithm, and MAC algorithm), securely exchanges the key between the server and client, and implements identity authentication of the server and client. Through the SSL handshake protocol, a session is established between a client and the server. A session consists of a set of parameters, including the session ID, peer certificate, cipher suite, and master secret.
- **SSL change cipher spec protocol**—Used for notification between the client and the server that the subsequent packets are to be protected and transmitted based on the newly negotiated cipher suite and key.
- **SSL alert protocol**—Enables the SSL client and server to send alert messages to each other. An alert message contains the alert severity level and a description.

## Configuration task list

| Task                                             | Remarks  |
|--------------------------------------------------|----------|
| <a href="#">Configuring an SSL server policy</a> | Required |
| <a href="#">Configuring an SSL client policy</a> | Optional |

## Configuring an SSL server policy

An SSL server policy is a set of SSL parameters for a server to use when booting up. An SSL server policy takes effect only after it is associated with an application layer protocol such as HTTP.

Before configuring an SSL server policy, configure the PKI domain for the SSL server policy to use to obtain a certificate for the SSL server. For more information about PKI domain configuration, see "[Configuring PKI.](#)"

SSL mainly comes in these versions: SSL 2.0, SSL 3.0, and TLS 1.0, where TLS 1.0 corresponds to SSL 3.1. When the switch acts as an SSL server, it can communicate with clients running SSL 3.0 or TLS 1.0, and can identify the SSL 2.0 Client Hello message from a client supporting SSL 2.0 and SSL 3.0/TLS 1.0 and notify the client to use SSL 3.0 or TLS 1.0 to communicate with the server.

To configure an SSL server policy:

| Step                  | Command                  | Remarks |
|-----------------------|--------------------------|---------|
| 1. Enter system view. | <code>system-view</code> | N/A     |

| Step                                                                                          | Command                                                                                                                                                                                      | Remarks                                                                                                                                                                                                                                                                                 |
|-----------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2. Create an SSL server policy and enter its view.                                            | <b>ssl server-policy</b> <i>policy-name</i>                                                                                                                                                  | N/A                                                                                                                                                                                                                                                                                     |
| 3. Specify a PKI domain for the SSL server policy.                                            | <b>pki-domain</b> <i>domain-name</i>                                                                                                                                                         | By default, no PKI domain is specified for an SSL server policy.<br>If the client requires certificate-based authentication for the SSL server, you must use this command to specify a PKI domain for the server and request a local certificate for the server through the PKI domain. |
| 4. Specify the cipher suite(s) for the SSL server policy to support.                          | <b>ciphersuite</b><br>[ <i>rsa_3des_edc_cbc_sha</i>   <i>rsa_aes_128_cbc_sha</i>   <i>rsa_aes_256_cbc_sha</i>   <i>rsa_des_cbc_sha</i>   <i>rsa_rc4_128_md5</i>   <i>rsa_rc4_128_sha</i> ] * | Optional.<br>By default, an SSL server policy supports all cipher suites.                                                                                                                                                                                                               |
| 5. Set the handshake timeout time for the SSL server.                                         | <b>handshake timeout</b> <i>time</i>                                                                                                                                                         | Optional.<br>3,600 seconds by default.                                                                                                                                                                                                                                                  |
| 6. Set the SSL connection close mode.                                                         | <b>close-mode</b> <i>wait</i>                                                                                                                                                                | Optional.<br>Not wait by default.                                                                                                                                                                                                                                                       |
| 7. Set the maximum number of cached sessions and the caching timeout time.                    | <b>session</b> { <i>cache-size size</i>   <i>timeout time</i> } *                                                                                                                            | Optional.<br>The defaults are as follows: <ul style="list-style-type: none"> <li>• 500 for the maximum number of cached sessions.</li> <li>• 3600 seconds for the caching timeout time.</li> </ul>                                                                                      |
| 8. Enable the SSL server to perform digital certificate-based authentication for SSL clients. | <b>client-verify</b> <i>enable</i>                                                                                                                                                           | Optional.<br>By default, the SSL server does not require clients to be authenticated.                                                                                                                                                                                                   |
| 9. Enable SSL client weak authentication.                                                     | <b>client-verify</b> <i>weaken</i>                                                                                                                                                           | Optional.<br>Disabled by default.<br>This command takes effect only when the <b>client-verify enable</b> command is configured.                                                                                                                                                         |

## SSL server policy configuration example

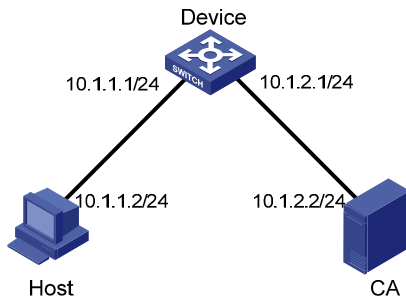
### Network requirements

As shown in [Figure 118](#), users need to access and control the device through web pages.

For security of the device and to make sure that data is not eavesdropped or tampered with, configure the device so that users must use HTTPS (Hypertext Transfer Protocol Secure, which uses SSL) to log in to the web interface of the device.



Figure 118 Network diagram



## Configuration considerations

To achieve the goal, perform the following configurations:

- Configure Device to work as the HTTPS server and request a certificate for Device.
- Request a certificate for Host so that Device can authenticate the identity of Host.
- Configure a CA server to issue certificates to Device and Host.

## Configuration procedure

In this example, Windows Server works as the CA server and the Simple Certificate Enrollment Protocol (SCEP) plug-in is installed on the CA server.

Before performing the following configurations, make sure the switch, the host, and the CA server can reach each other.

1. Configure the HTTPS server (Device):

# Create a PKI entity named **en**, and configure the common name as **http-server1** and the FQDN as **ssl.security.com**.

```
<Device> system-view
[Device] pki entity en
[Device-pki-entity-en] common-name http-server1
[Device-pki-entity-en] fqdn ssl.security.com
[Device-pki-entity-en] quit
```

# Create PKI domain **1**, specify the trusted CA as **ca server**, the URL of the registration server as **http://10.1.2.2/certsrv/mscep/mscep.dll**, the authority for certificate request as RA, and the entity for certificate request as **en**.

```
[Device] pki domain 1
[Device-pki-domain-1] ca identifier ca server
[Device-pki-domain-1] certificate request url
http://10.1.2.2/certsrv/mscep/mscep.dll
[Device-pki-domain-1] certificate request from ra
[Device-pki-domain-1] certificate request entity en
[Device-pki-domain-1] quit
```

# Create the local RSA key pairs.

```
[Device] public-key local create rsa
```

# Retrieve the CA certificate.

```
[Device] pki retrieval-certificate ca domain 1
```

# Request a local certificate for Device.

```
[Device] pki request-certificate domain 1
```

```

Create an SSL server policy named myssl.
[Device] ssl server-policy myssl
Specify the PKI domain for the SSL server policy as 1.
[Device-ssl-server-policy-myssl] pki-domain 1
Enable client authentication.
[Device-ssl-server-policy-myssl] client-verify enable
[Device-ssl-server-policy-myssl] quit
Configure HTTPS service to use SSL server policy myssl.
[Device] ip https ssl-server-policy myssl
Enable HTTPS service.
[Device] ip https enable
Create a local user named usera, and set the password to 123 and service type to telnet.
[Device] local-user usera
[Device-luser-usera] password simple 123
[Device-luser-usera] service-type telnet

```

**2.** Configure the HTTPS client (Host):

On Host, launch IE, enter `http://10.1.2.2/certsrv` in the address bar and request a certificate for Host as prompted.

**3.** Verify your configuration:

Launch IE on the host, enter `https://10.1.1.1` in the address bar, and select the certificate issued by the CA server. The web interface of the switch should appear. After entering username **usera** and password **123**, you should be able to log in to the web interface to access and manage the switch.

For more information about PKI configuration commands, see "[Configuring PKI](#)."

For more information about the **public-key local create rsa** command, see *Security Command Reference*.

For more information about HTTPS, see *Fundamentals Configuration Guide*.

## Configuring an SSL client policy

An SSL client policy is a set of SSL parameters for a client to use when connecting to the server. An SSL client policy takes effect only after it is associated with an application layer protocol.

If the SSL server is configured to authenticate the SSL client, you must configure the PKI domain for the SSL client policy to use to obtain the certificate of the client. For more information about PKI domain configuration, see "[Configuring PKI](#)."

To configure an SSL client policy:

| Step                                                      | Command                                     | Remarks |
|-----------------------------------------------------------|---------------------------------------------|---------|
| <b>1.</b> Enter system view.                              | <b>system-view</b>                          | N/A     |
| <b>2.</b> Create an SSL client policy and enter its view. | <b>ssl client-policy</b> <i>policy-name</i> | N/A     |

| Step                                                                                     | Command                                                                                                                                                                                                     | Remarks                                                                                                                                                                                                                                                                      |
|------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 3. Specify a PKI domain for the SSL client policy.                                       | <b>pki-domain</b> <i>domain-name</i>                                                                                                                                                                        | Optional.<br>No PKI domain is configured by default.<br>If the SSL server requires certificate-based authentication for SSL clients, you must use this command to specify a PKI domain for the client and request a local certificate for the client through the PKI domain. |
| 4. Specify the preferred cipher suite for the SSL client policy.                         | <b>prefer-cipher</b><br>{ <i>rsa_3des_edc_cbc_sha</i>  <br><i>rsa_aes_128_cbc_sha</i>  <br><i>rsa_aes_256_cbc_sha</i>  <br><i>rsa_des_cbc_sha</i>  <br><i>rsa_rc4_128_md5</i>  <br><i>rsa_rc4_128_sha</i> } | Optional.<br><b>rsa_rc4_128_md5</b> by default.                                                                                                                                                                                                                              |
| 5. Specify the SSL protocol version for the SSL client policy.                           | <b>version</b> { <i>ssl3.0</i>   <i>tls1.0</i> }                                                                                                                                                            | Optional.<br>TLS 1.0 by default.                                                                                                                                                                                                                                             |
| 6. Enable the SSL client to perform certificate-based authentication for the SSL server. | <b>server-verify</b> <b>enable</b>                                                                                                                                                                          | Optional.<br>Enabled by default.                                                                                                                                                                                                                                             |

## Displaying and maintaining SSL

| Task                                   | Command                                                                                                                                                          | Remarks               |
|----------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Display SSL server policy information. | <b>display ssl server-policy</b><br>{ <i>policy-name</i>   <b>all</b> } [   { <b>begin</b>  <br><b>exclude</b>   <b>include</b> }<br><i>regular-expression</i> ] | Available in any view |
| Display SSL client policy information. | <b>display ssl client-policy</b><br>{ <i>policy-name</i>   <b>all</b> } [   { <b>begin</b>  <br><b>exclude</b>   <b>include</b> }<br><i>regular-expression</i> ] | Available in any view |

## Troubleshooting SSL

### Symptom

As the SSL server, the switch fails to handshake with the SSL client.

### Analysis

SSL handshake failure may result from the following causes:

- The SSL client is configured to authenticate the SSL server, but the SSL server has no certificate or the certificate is not trusted.

- The SSL server is configured to authenticate the SSL client, but the SSL client has no certificate or the certificate is not trusted.
- The server and the client have no matching cipher suite.

## Solution

1. Issue the **debugging ssl** command and view the debugging information to locate the problem:
  - If the SSL client is configured to authenticate the SSL server but the SSL server has no certificate, request one for it.
  - If the server's certificate cannot be trusted, install the root certificate of the CA that issued the local certificate to the SSL server on the SSL client, or let the server request a certificate from the CA that the SSL client trusts.
  - If the SSL server is configured to authenticate the client, but the SSL client has no certificate or the certificate cannot be trusted, request and install a certificate for the client.
2. Use the **display ssl server-policy** command to view the cipher suites that the SSL server policy supports. If the server and the client have no matching cipher suite, use the **ciphersuite** command to modify the cipher suite configuration of the SSL server.

---

# Configuring TCP attack protection

## Overview

An attacker can attack the switch during the process of establishing a TCP connection. To prevent such an attack, the switch provides the SYN Cookie feature.

## Enabling the SYN Cookie feature

As a general rule, the establishment of a TCP connection involves the following three handshakes.

1. The request originator sends a SYN message to the target server.
2. After receiving the SYN message, the target server establishes a TCP connection in the SYN\_RECEIVED state, returns a SYN ACK message to the originator, and waits for a response.
3. After receiving the SYN ACK message, the originator returns an ACK message, establishing the TCP connection.

Attackers may mount SYN Flood attacks during TCP connection establishment. They send a large number of SYN messages to the server to establish TCP connections, but they never make any response to SYN ACK messages. As a result, a large number of incomplete TCP connections are established, resulting in heavy resource consumption and making the server unable to handle services normally.

The SYN Cookie feature can prevent SYN Flood attacks. After receiving a TCP connection request, the server directly returns a SYN ACK message, instead of establishing an incomplete TCP connection. Only after receiving an ACK message from the client can the server establish a connection, and then enter the ESTABLISHED state. In this way, incomplete TCP connections could be avoided to protect the server against SYN Flood attacks.

Follow these guidelines when you enable the SYN Cookie feature:

- If you enable MD5 authentication for TCP connections on the HP 5500 EI, the SYN Cookie configuration is ineffective. Then, if you disable MD5 authentication for TCP connections, the SYN Cookie configuration automatically becomes effective. For more information about MD5 authentication, see *Layer 3—IP Routing Configuration Guide*.
- With the SYN Cookie feature enabled, only the maximum segment size (MSS), is negotiated during TCP connection establishment, instead of the window's zoom factor and timestamp.

To enable the SYN Cookie feature:

| Step                              | Command                      | Remarks            |
|-----------------------------------|------------------------------|--------------------|
| 1. Enter system view.             | <b>system-view</b>           | N/A                |
| 2. Enable the SYN Cookie feature. | <b>tcp syn-cookie enable</b> | Enabled by default |

## Displaying and maintaining TCP attack protection

| Task                                  | Command                                                                                                      | Remarks               |
|---------------------------------------|--------------------------------------------------------------------------------------------------------------|-----------------------|
| Display current TCP connection state. | <b>display tcp status</b> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] | Available in any view |

# Configuring IP source guard

## Overview

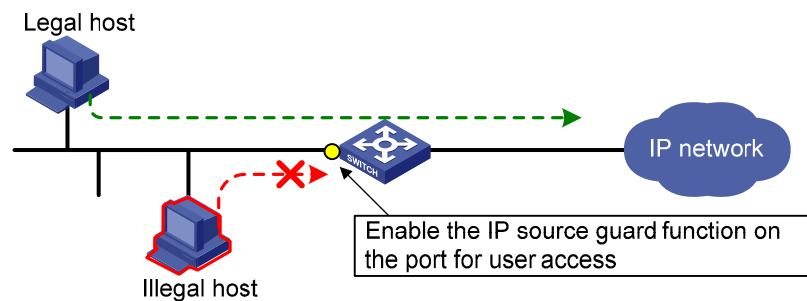
IP source guard is intended to improve port security by blocking illegal packets. For example, it can prevent illegal hosts from using a legal IP address to access the network.

IP source guard can filter packets according to the packet source IP address and source MAC address. IP source guard entries fall into the following types:

- IP-port binding entry
- MAC-port binding entry
- IP-MAC-port binding entry

After receiving a packet, an IP source guard-enabled port obtains the key attributes (source IP address and source MAC) of the packet and then looks them up in the IP source guard entries. If there is a match, the port forwards the packet. Otherwise, the port discards the packet, as shown in [Figure 119](#).

**Figure 119 Diagram for the IP source guard function**



A binding entry can be statically configured or dynamically added.

## Static IP source guard entries

A static IP source guard entry is configured manually. It is suitable for scenarios where few hosts exist on a LAN and their IP addresses are manually configured. For example, you can configure a static binding entry on a port that connects a server, allowing the port to receive packets from and send packets to only the server.

A static IPv4 source guard entry filters IPv4 packets received by the port or cooperates with ARP detection to check the validity of users. A static IPv6 source guard entry filters IPv6 packets received by the port cooperates with the ND detection feature to check the validity of users.

For information about ARP detection, see "[Configuring ARP attack protection](#)." For information about ND detection, see "[Configuring ND attack defense](#)."

A static IP source guard entry can be a global or port-based static binding entry.

## Global static binding entry

A global static binding entry is a MAC-IP binding entry configured in system view. It is effective on all ports. A port forwards a packet when the packet's IP address and MAC address both match those of a global static binding entry or a static binding entry configured on the port.

Global static binding entries are used to protect against host spoofing attacks, which exploit the IP address or MAC address of a legal user host.

## Port-based static binding entry

A port-based static binding entry binds an IP address, MAC address, or any combination of the three with a port. Such an entry is effective on only the specified port. A port forwards a packet only when the IP address and MAC address (if any) of the packet all match those in a static binding entry on the port or a global static binding entry. All other packets will be dropped.

Port-based static binding entries are used to check the validity of users who are trying to access a port.

## Dynamic IP source guard entries

Dynamic IP source guard entries are generated dynamically according to client entries on the DHCP snooping or DHCP relay agent device. They are suitable for scenarios where many hosts reside on a LAN and obtain IP addresses through DHCP. Once DHCP allocates an IP address to a client, IP source guard automatically adds the client entry to allow the client to access the network. A user using an IP address not obtained through DHCP cannot access the network. Dynamic IPv6 source guard entries can also be obtained from client entries on the ND snooping device.

- Dynamic IPv4 source guard entries are generated dynamically based on DHCP snooping or DHCP relay entries to filter incoming IPv4 packets on a port.
- Dynamic IPv6 source guard entries are generated dynamically based on DHCPv6 snooping or ND snooping entries to filter incoming IPv6 packets on a port.

For information about DHCP snooping, DHCP relay, DHCPv6 snooping, and ND snooping, see *Layer 3—IP Services Configuration Guide*.

## Configuration task list

Complete the following tasks to configure IPv4 source guard:

| Task                                                                    | Remarks  |
|-------------------------------------------------------------------------|----------|
| <a href="#">Configuring IPv4 source guard on a port</a>                 | Required |
| <a href="#">Configuring a static IPv4 source guard entry</a>            | Optional |
| <a href="#">Setting the maximum number of IPv4 source guard entries</a> | Optional |

Complete the following tasks to configure IPv6 source guard:

| Task                                                                    | Remarks  |
|-------------------------------------------------------------------------|----------|
| <a href="#">Configuring IPv6 source guard on a port</a>                 | Required |
| <a href="#">Configuring a static IPv6 source guard entry</a>            | Optional |
| <a href="#">Setting the maximum number of IPv6 source guard entries</a> | Optional |



# Configuring the IPv4 source guard function

You cannot enable IPv4 source guard on a link aggregation member port or a service loopback group. If IPv4 source guard is enabled on a port, you cannot assign the port to a link aggregation group or a service loopback group.

## Configuring IPv4 source guard on a port

The IPv4 source guard function must be configured on a port before the port can obtain dynamic IPv4 source guard entries and use static and dynamic IPv4 source guard entries to filter packets.

- For how to configure a static binding entry, see "[Configuring a static IPv4 source guard entry.](#)"
- On a Layer 2 Ethernet port, IP source guard cooperates with DHCP snooping, dynamically obtains the DHCP snooping entries generated during dynamic IP address allocation, and generates IP source guard entries accordingly.
- On a VLAN interface, IP source guard cooperates with DHCP relay, dynamically obtains the DHCP relay entries generated during dynamic IP address allocation across network segments, and generates IP source guard entries accordingly.

Dynamic IPv4 source guard entries can contain such information as the MAC address, IP address, VLAN tag, ingress port information, and entry type (DHCP snooping or DHCP relay), where the MAC address, IP address, or VLAN tag information may not be included depending on your configuration. IP source guard applies these entries to the port to filter packets.

To generate IPv4 binding entries dynamically based on DHCP entries, make sure that DHCP snooping or DHCP relay is configured and working normally. For information about DHCP snooping configuration and DHCP relay configuration, see *Layer 3—IP Services Configuration Guide*.

If you repeatedly configure the IPv4 source guard function on a port, only the last configuration takes effect.

To configure the IPv4 source guard function on a port:

| Step                                        | Command                                                                                                  | Remarks                                                                                                                                                               |
|---------------------------------------------|----------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                       | <b>system-view</b>                                                                                       | N/A                                                                                                                                                                   |
| 2. Enter interface view.                    | <b>interface</b> <i>interface-type</i><br><i>interface-number</i>                                        | The term <i>interface</i> collectively refers to the following types of ports and interfaces: Bridge mode (Layer 2) Ethernet ports, VLAN interfaces, and port groups. |
| 3. Configure IPv4 source guard on the port. | <b>ip verify source</b> { <b>ip-address</b>  <br><b>ip-address mac-address</b>  <br><b>mac-address</b> } | Not configured by default.                                                                                                                                            |

### NOTE:

Although dynamic IPv4 source guard entries are generated based on DHCP entries, the number of dynamic IPv4 source guard entries is not necessarily the same as that of the DHCP entries.

# Configuring a static IPv4 source guard entry

Static IPv4 binding entries take effect only on the ports configured with the IPv4 source guard function (see "Configuring IPv4 source guard on a port").

Port-based static IPv4 source guard entries and dynamic IPv4 source guard entries take precedence over global static IPv4 source guard entries. A port matches a packet against global static binding entries only when the packet does not match any port-based static binding entry or dynamic binding entry on the port.

## Configuring global static IPv4 binding entries

A global static binding entry defines the IP address and MAC address of the packets that can be forwarded by ports. It takes effect on all ports of the device.

To configure a global static IPv4 binding entry:

| Step                                             | Command                                                                | Remarks                                                       |
|--------------------------------------------------|------------------------------------------------------------------------|---------------------------------------------------------------|
| 1. Enter system view.                            | <b>system-view</b>                                                     | N/A                                                           |
| 2. Configure a global static IPv4 binding entry. | <b>ip source binding ip-address ip-address mac-address mac-address</b> | No global static IPv4 binding entry is configured by default. |

## Configuring port-based static IPv4 binding entries

Follow these guidelines to configure port-based static IPv4 source guard entries:

- You cannot repeatedly configure the same static binding entry on one port, but you can configure the same static entry on different ports.
- IP source guard does not use the VLAN information (if specified) in static IPv4 binding entries to filter packets.
- When the ARP detection function is configured, be sure to specify the VLAN where ARP detection is configured in static IPv4 binding entries. Otherwise, ARP packets are discarded because they cannot match any static IPv4 binding entry.
- If a static binding entry to be added denotes the same binding as an existing dynamic binding entry, the new static binding entry overwrites the dynamic binding entry.

To configure a static IPv4 binding entry on a port:

| Step                                                       | Command                                                                                                                                       | Remarks                                                           |
|------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------|
| 1. Enter system view.                                      | <b>system-view</b>                                                                                                                            | N/A                                                               |
| 2. Enter Layer 2 interface view.                           | <b>interface interface-type interface-number</b>                                                                                              | N/A                                                               |
| 3. Configure a static IPv4 source guard entry on the port. | <b>ip source binding { ip-address ip-address   ip-address ip-address mac-address mac-address   mac-address mac-address } [ vlan vlan-id ]</b> | By default, no static IPv4 binding entry is configured on a port. |

## Setting the maximum number of IPv4 source guard entries

The maximum number of IPv4 source guard entries is used to limit the total number of static and dynamic IPv4 source guard entries on a port. When the number of IPv4 binding entries on a port reaches the maximum, the port does not allowed new IPv4 binding entries any more.

If the maximum number of IPv4 binding entries to be configured is smaller than the number of existing IPv4 binding entries on the port, the maximum number can be configured successfully, and the existing entries are not affected. New IPv4 binding entries, however, cannot be added until the number of IPv4 binding entries on the port drops below the configured maximum.

To configure the maximum number of IPv4 binding entries allowed on a port:

| Step                                                                         | Command                                                           | Remarks                                                                                                        |
|------------------------------------------------------------------------------|-------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                                                        | <b>system-view</b>                                                | N/A                                                                                                            |
| 2. Enter Layer 2 Ethernet interface view.                                    | <b>interface</b> <i>interface-type</i><br><i>interface-number</i> | N/A                                                                                                            |
| 3. Configure the maximum number of IPv4 binding entries allowed on the port. | <b>ip verify source max-entries</b><br><i>number</i>              | Optional.<br>By default, the maximum number is 1500 on the HP 5500 EI series and 640 on the HP 5500 SI series. |

## Configuring the IPv6 source guard function

You cannot enable IPv6 source guard on a link aggregation member port or a service loopback port. If IPv6 source guard is enabled on a port, you cannot assign the port to a link aggregation group or a service loopback group.

### Configuring IPv6 source guard on a port

The IPv6 source guard function must be configured on a port before the port can obtain dynamic IPv6 source guard entries and use static and dynamic IPv6 source guard entries to filter packets.

- For how to configure a static IPv6 static binding entry, see "[Configuring a static IPv6 source guard entry.](#)"
- Cooperating with DHCPv6 snooping, IP source guard dynamically generates IP source guard entries based on the DHCPv6 snooping entries that are generated during dynamic IP address allocation.
- Cooperating with ND snooping, IP source guard dynamically generates IP source guard entries based on dynamic ND snooping entries.

Dynamic IPv6 source guard entries can contain such information as the MAC address, IPv6 address, VLAN tag, ingress port information and entry type (DHCPv6 snooping or ND snooping), where the MAC address, IPv6 address, and/or VLAN tag information may not be included depending on your configuration. IP source guard applies these entries to the port, so that the port can filter packets accordingly.

Follow these guidelines when you configure IPv6 source guard:

- If you repeatedly configure the IPv6 source guard function, only the last configuration takes effect.

- To obtain dynamic IPv6 source guard entries, make sure that DHCPv6 snooping or ND snooping is configured and works normally. For DHCPv6 and ND snooping configuration information, see *Layer 3—IP Services Configuration Guide*.
- If you configure both ND snooping and DHCPv6 snooping on the device, IPv6 source guard uses the type of entries that generated first. Because DHCPv6 snooping entries are usually generated first in such a case, IPv6 source guard usually uses the DHCPv6 snooping entries to filter packets on a port.

To configure the IPv6 source guard function on a port:

| Step                                                       | Command                                                                                                  | Remarks                                                                                                                                                                                                                                                                                                          |
|------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                                      | <b>system-view</b>                                                                                       | N/A                                                                                                                                                                                                                                                                                                              |
| 2. Enter Layer 2 Ethernet interface view, port group view. | <b>interface</b> <i>interface-type</i><br><i>interface-number</i>                                        | N/A                                                                                                                                                                                                                                                                                                              |
| 3. Configure the IPv6 source guard function on the port.   | <b>ipv6 verify source</b> { <b>ipv6-address</b>   <b>ipv6-address mac-address</b>   <b>mac-address</b> } | Not configured by default.<br>The keyword specified in the <b>ipv6 verify source</b> command is only for instructing the generation of dynamic IPv6 source guard entries. It does not affect static binding entries. When using a static binding entry, a port does not consider the keyword into consideration. |

#### NOTE:

Although dynamic IPv6 source guard entries are generated based on DHCPv6 entries, the number of dynamic IPv6 source guard entries is not necessarily the same as that of the DHCPv6 entries.

## Configuring a static IPv6 source guard entry

Static IPv6 binding entries take effect only on ports configured with the IPv6 source guard function (see "[Configuring the IPv6 source guard function](#)").

Port-based static IPv6 source guard entries and dynamic IPv6 source guard entries take precedence over global static IPv6 source guard entries. A port matches a packet against global static binding entries only when the packet does not match any port-based static binding entry or dynamic binding entry on the port.

### Configuring global static IPv6 binding entries

A global static IPv6 binding entry defines the IPv6 address and MAC address of the packets that can be forwarded by ports. It takes effect on all ports of the device.

To configure a global static IPv6 binding entry:

| Step                  | Command            | Remarks |
|-----------------------|--------------------|---------|
| 1. Enter system view. | <b>system-view</b> | N/A     |

| Step                                             | Command                                                                                          | Remarks                                                       |
|--------------------------------------------------|--------------------------------------------------------------------------------------------------|---------------------------------------------------------------|
| 2. Configure a global static IPv6 binding entry. | <b>ipv6 source binding ipv6-address</b><br><i>ipv6-address mac-address</i><br><i>mac-address</i> | No global static IPv6 binding entry is configured by default. |

## Configuring port-based static IPv6 binding entries

Follow these guidelines to configure port-based static IPv6 source guard entries:

- You cannot configure the same static binding entry on one port repeatedly, but you can configure the same static binding entry on different ports.
- In an IPv6 source guard entry, the MAC address cannot be all 0s, all Fs (a broadcast MAC address), or a multicast address, and the IPv6 address must be a unicast address and cannot be all 0s, all Fs, or a loopback address.
- IP source guard does not use the VLAN information (if specified) in static IPv6 binding entries to filter packets.
- When the ND detection function is configured, be sure to specify the VLAN where ND detection is configured in static binding entries. Otherwise, ND packets will be discarded because they cannot match any static IPv6 binding entry.
- If a static binding entry to be added denotes the same binding as an existing dynamic binding entry, the new static binding entry overwrites the dynamic binding entry.

To configure a static IPv6 source guard entry on a port:

| Step                                                | Command                                                                                                                                                                                                                     | Remarks                                                           |
|-----------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------|
| 1. Enter system view.                               | <b>system-view</b>                                                                                                                                                                                                          | N/A                                                               |
| 2. Enter Layer 2 interface view.                    | <b>interface</b> <i>interface-type</i><br><i>interface-number</i>                                                                                                                                                           | N/A                                                               |
| 3. Configure a static IPv6 binding entry on a port. | <b>ipv6 source binding { ipv6-address</b><br><i>ipv6-address</i>   <b>ipv6-address</b><br><i>ipv6-address mac-address</i><br><i>mac-address</i>   <b>mac-address</b><br><i>mac-address</i> } [ <b>vlan</b> <i>vlan-id</i> ] | By default, no static IPv6 binding entry is configured on a port. |

## Setting the maximum number of IPv6 source guard entries

The maximum number of IPv6 source guard entries is used to limit the total number of static and dynamic IPv6 source guard entries on a port. When the number of IPv6 binding entries on a port reaches the maximum, the port does not allow new IPv6 binding entries any more.

If the maximum number of IPv6 binding entries to be configured is smaller than the number of existing IPv6 binding entries on the port, the maximum number can be configured successfully, and the existing entries are not affected. New IPv6 binding entries, however, cannot be added until the number of IPv6 binding entries on the port drops below the configured maximum.

To configure the maximum number of IPv6 binding entries allowed on a port:

| Step                  | Command            | Remarks |
|-----------------------|--------------------|---------|
| 1. Enter system view. | <b>system-view</b> | N/A     |

| Step                                                                         | Command                                                           | Remarks                                                                                                        |
|------------------------------------------------------------------------------|-------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------|
| 2. Enter Layer 2 Ethernet interface view.                                    | <b>interface</b> <i>interface-type</i><br><i>interface-number</i> | N/A                                                                                                            |
| 3. Configure the maximum number of IPv6 binding entries allowed on the port. | <b>ipv6 verify source max-entries</b><br><i>number</i>            | Optional.<br>By default, the maximum number is 1500 on the HP 5500 EI series and 640 on the HP 5500 SI series. |

## Displaying and maintaining IP source guard

For IPv4 source guard:

| Task                                      | Command                                                                                                                                                                                                                                                                                                             | Remarks               |
|-------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Display static IPv4 source guard entries. | <b>display ip source binding static</b> [ <b>interface</b><br><i>interface-type interface-number</i>  <br><b>ip-address</b> <i>ip-address</i>   <b>mac-address</b><br><i>mac-address</i> ] [ <b>slot</b> <i>slot-number</i> ] [   { <b>begin</b>  <br><b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] | Available in any view |
| Display IPv4 source guard entries.        | <b>display ip source binding</b> [ <b>interface</b><br><i>interface-type interface-number</i>  <br><b>ip-address</b> <i>ip-address</i>   <b>mac-address</b><br><i>mac-address</i> ] [ <b>slot</b> <i>slot-number</i> ] [   { <b>begin</b>  <br><b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]        | Available in any view |

For IPv6 source guard:

| Task                                      | Command                                                                                                                                                                                                                                                                                                                   | Remarks               |
|-------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Display static IPv6 source guard entries. | <b>display ipv6 source binding static</b> [ <b>interface</b><br><i>interface-type interface-number</i>  <br><b>ipv6-address</b> <i>ipv6-address</i>   <b>mac-address</b><br><i>mac-address</i> ] [ <b>slot</b> <i>slot-number</i> ] [   { <b>begin</b>  <br><b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] | Available in any view |
| Display IPv6 source guard entries.        | <b>display ipv6 source binding</b> [ <b>interface</b><br><i>interface-type interface-number</i>  <br><b>ipv6-address</b> <i>ipv6-address</i>   <b>mac-address</b><br><i>mac-address</i> ] [ <b>slot</b> <i>slot-number</i> ] [   { <b>begin</b>  <br><b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]        | Available in any view |

## IP source guard configuration examples

### Static IPv4 source guard configuration example

#### Network requirements

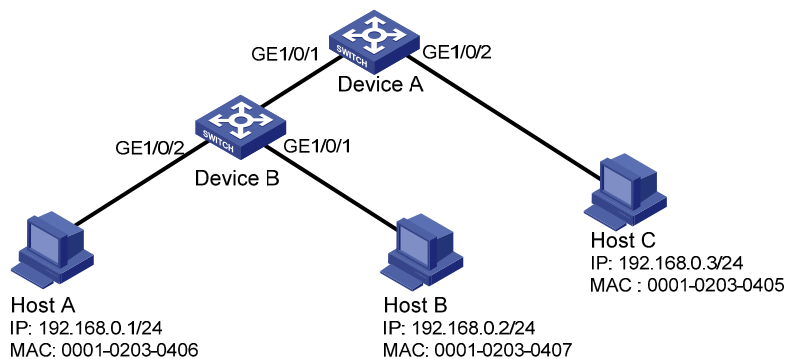
As shown in [Figure 120](#), Host A and Host B are connected to ports GigabitEthernet 1/0/2 and GigabitEthernet 1/0/1 of Device B respectively, Host C is connected to port GigabitEthernet 1/0/2 of

Device A, and Device B is connected to port GigabitEthernet 1/0/1 of Device A. All hosts use static IP addresses.

Configure static IPv4 source guard entries on Device A and Device B to meet the following requirements:

- On port GigabitEthernet 1/0/2 of Device A, only IP packets from Host C can pass.
- On port GigabitEthernet 1/0/1 of Device A, only IP packets from Host A can pass.
- On port GigabitEthernet 1/0/2 of Device B, only IP packets from Host A can pass.
- On port GigabitEthernet 1/0/1 of Device B, only IP packets sourced from 192.168.0.2/24 can pass. Host B can communicate with Host A by using this IP address even if it uses another network adapter.

**Figure 120 Network diagram**



## Configuration procedure

### 1. Configure Device A:

# Configure the IPv4 source guard function on GigabitEthernet 1/0/2 to filter packets based on both the source IP address and MAC address.

```
<DeviceA> system-view
```

```
[DeviceA] interface gigabitethernet 1/0/2
```

```
[DeviceA-GigabitEthernet1/0/2] ip verify source ip-address mac-address
```

# Configure GigabitEthernet 1/0/2 to allow only IP packets with the source MAC address of 0001-0203-0405 and the source IP address of 192.168.0.3 to pass.

```
[DeviceA] interface gigabitethernet 1/0/2
```

```
[DeviceA-GigabitEthernet1/0/2] ip source binding ip-address 192.168.0.3 mac-address 0001-0203-0405
```

```
[DeviceA-GigabitEthernet1/0/2] quit
```

# Configure the IPv4 source guard function on GigabitEthernet 1/0/1 to filter packets based on both the source IP address and MAC address.

```
[DeviceA] interface gigabitethernet 1/0/1
```

```
[DeviceA-GigabitEthernet1/0/1] ip verify source ip-address mac-address
```

# Configure GigabitEthernet 1/0/1 to allow only IP packets with the source MAC address of 0001-0203-0406 and the source IP address of 192.168.0.1 to pass.

```
[DeviceA-GigabitEthernet1/0/1] ip source binding ip-address 192.168.0.1 mac-address 0001-0203-0406
```

```
[DeviceA-GigabitEthernet1/0/1] quit
```

### 2. Configure Device B:

# Configure the IPv4 source guard function on GigabitEthernet 1/0/2 to filter packets based on both the source IP address and MAC address.

```
[DeviceB] interface gigabitethernet 1/0/2
```

```
[DeviceB-GigabitEthernet1/0/2] ip verify source ip-address mac-address
```

# Configure GigabitEthernet 1/0/2 to allow only IP packets with the source MAC address of 0001-0203-0406 and the source IP address of 192.168.0.1 to pass.

```
[DeviceB] interface gigabitethernet 1/0/2
```

```
[DeviceB-GigabitEthernet1/0/2] ip source binding ip-address 192.168.0.1 mac-address 0001-0203-0406
```

```
[DeviceB-GigabitEthernet1/0/2] quit
```

# Configure the IPv4 source guard function on GigabitEthernet 1/0/1 to filter packets based on the source IP address.

```
[DeviceB] interface gigabitethernet 1/0/1
```

```
[DeviceB-GigabitEthernet1/0/1] ip verify source ip-address
```

# Configure GigabitEthernet 1/0/1 to allow only IP packets with the source IP address of 192.168.0.2 to pass.

```
[DeviceB-GigabitEthernet1/0/1] ip source binding ip-address 192.168.0.2
```

```
[DeviceB-GigabitEthernet1/0/1] quit
```

## Verifying the configuration

# On Device A, display information about static IPv4 source guard entries. The output shows that the static IPv4 source guard entries are configured successfully.

```
[DeviceA] display ip source binding static
```

```
Total entries found: 2
```

| MAC Address    | IP Address  | VLAN | Interface | Type   |
|----------------|-------------|------|-----------|--------|
| 0001-0203-0405 | 192.168.0.3 | N/A  | GE1/0/2   | Static |
| 0001-0203-0406 | 192.168.0.1 | N/A  | GE1/0/1   | Static |

# On Device B, display information about static IPv4 source guard entries. The output shows that the static IPv4 source guard entries are configured successfully.

```
[DeviceB] display ip source binding static
```

```
Total entries found: 2
```

| MAC Address    | IP Address  | VLAN | Interface | Type   |
|----------------|-------------|------|-----------|--------|
| 0001-0203-0406 | 192.168.0.1 | N/A  | GE1/0/2   | Static |
| N/A            | 192.168.0.2 | N/A  | GE1/0/1   | Static |

## Dynamic IPv4 source guard using DHCP snooping configuration example

### Network requirements

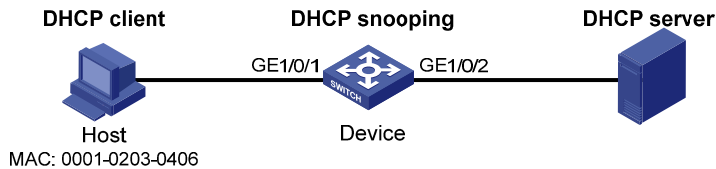
As shown in [Figure 121](#), the device connects to the host (client) and the DHCP server through ports GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 respectively. The host obtains an IP address from the DHCP server.

Enable DHCP snooping on the device to record the DHCP snooping entry of the host. Enable the IPv4 source guard function on the device's port GigabitEthernet 1/0/1 to filter packets based on the DHCP snooping entry, allowing only packets from clients that obtain IP addresses through the DHCP server to pass.



For information about DHCP server configuration, see *Layer 3—IP Services Configuration Guide*.

**Figure 121 Network diagram**



## Configuration procedure

1. Configure DHCP snooping.

# Enable DHCP snooping.

```
<Device> system-view
```

```
[Device] dhcp-snooping
```

# Configure port GigabitEthernet 1/0/2, which is connected to the DHCP server, as a trusted port.

```
[Device] interface gigabitethernet1/0/2
```

```
[Device-GigabitEthernet1/0/2] dhcp-snooping trust
```

```
[Device-GigabitEthernet1/0/2] quit
```

2. Configure the IPv4 source guard function.

# Configure the IPv4 source guard function on port GigabitEthernet 1/0/1 to filter packets based on both the source IP address and MAC address.

```
[Device] interface gigabitethernet1/0/1
```

```
[Device-GigabitEthernet1/0/1] ip verify source ip-address mac-address
```

```
[Device-GigabitEthernet1/0/1] quit
```

## Verifying the configuration

# Display the IPv4 source guard entries generated on port GigabitEthernet 1/0/1.

```
[Device] display ip source binding
```

```
Total entries found: 1
```

| MAC Address    | IP Address  | VLAN | Interface | Type     |
|----------------|-------------|------|-----------|----------|
| 0001-0203-0406 | 192.168.0.1 | 1    | GE1/0/1   | DHCP-SNP |

# Display DHCP snooping entries to see whether they are consistent with the dynamic entries generated on GigabitEthernet 1/0/1.

```
[Device] display dhcp-snooping
```

```
DHCP Snooping is enabled.
```

```
The client binding table for all untrusted ports.
```

```
Type : D--Dynamic , S--Static , R--Recovering
```

```
Type IP Address MAC Address Lease VLAN SVLAN Interface
```

```
==== =====
```

```
D 192.168.0.1 0001-0203-0406 86335 1 N/A GigabitEthernet1/0/1
```

```
--- 1 dhcp-snooping item(s) found ---
```

The output shows that a dynamic IPv4 source guard entry has been generated based on the DHCP snooping entry.

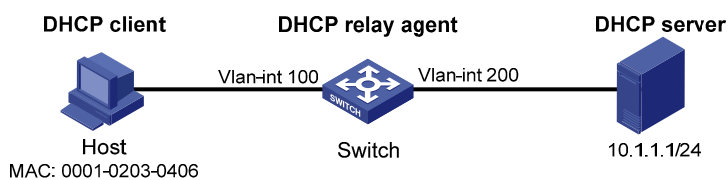
# Dynamic IPv4 source guard using DHCP relay configuration example

## Network requirements

As shown in Figure 122, the host and the DHCP server are connected to the switch through interfaces VLAN-interface 100 and VLAN-interface 200 respectively. DHCP relay is enabled on the switch. The host (with the MAC address of 0001-0203-0406) obtains an IP address from the DHCP server through the DHCP relay agent.

Enable the IPv4 source guard function on the switch's VLAN-interface 100 to filter packets based on the DHCP relay entry, allowing only packets from clients that obtain IP addresses from the DHCP server to pass.

Figure 122 Network diagram



## Configuration procedure

1. Configure the IPv4 source guard function:
  - # Configure the IP addresses of the interfaces. (Details not shown.)
  - # Configure the IPv4 source guard function on VLAN-interface 100 to filter packets based on both the source IP address and MAC address.

```
<Switch> system-view
[Switch] vlan 100
[Switch-Vlan100] quit
[Switch] interface vlan-interface 100
[Switch-Vlan-interface100] ip verify source ip-address mac-address
[Switch-Vlan-interface100] quit
```
2. Configure the DHCP relay agent:
  - # Enable the DHCP service.

```
[Switch] dhcp enable
```
  - # Configure the IP address of the DHCP server.

```
[Switch] dhcp relay server-group 1 ip 10.1.1.1
```
  - # Configure VLAN-interface 100 to operate in DHCP relay mode.

```
[Switch] interface vlan-interface 100
[Switch-Vlan-interface100] dhcp select relay
```
  - # Correlate VLAN-interface 100 with DHCP server group 1.

```
[Switch-Vlan-interface100] dhcp relay server-select 1
[Switch-Vlan-interface100] quit
```

## Verifying the configuration

- # Display the generated IPv4 source guard entries.

```
[Switch] display ip source binding
```

Total entries found: 1

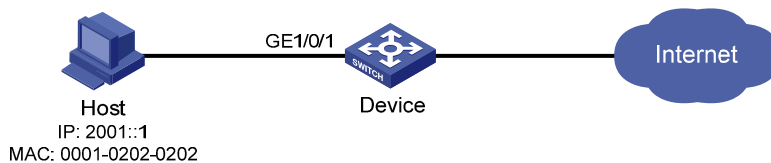
| MAC Address    | IP Address  | VLAN | Interface | Type     |
|----------------|-------------|------|-----------|----------|
| 0001-0203-0406 | 192.168.0.1 | 100  | Vlan100   | DHCP-RLY |

## Static IPv6 source guard configuration example

### Network requirements

As shown in [Figure 123](#), the host is connected to port GigabitEthernet 1/0/1 of the device. Configure a static IPv6 source guard entry for GigabitEthernet 1/0/1 of the device to allow only packets from the host to pass.

**Figure 123 Network diagram**



### Configuration procedure

# Configure the IPv6 source guard function on GigabitEthernet 1/0/1 to filter packets based on both the source IP address and MAC address.

```
<Device> system-view
[Device] interface gigabitethernet 1/0/1
[Device-GigabitEthernet1/0/1] ipv6 verify source ipv6-address mac-address
```

# Configure GigabitEthernet 1/0/1 to allow only IPv6 packets with the source MAC address of 0001-0202-0202 and the source IPv6 address of 2001::1 to pass.

```
[Device-GigabitEthernet1/0/1] ipv6 source binding ipv6-address 2001::1 mac-address
0001-0202-0202
[Device-GigabitEthernet1/0/1] quit
```

### Verifying the configuration

# On Device, display the information about static IPv6 source guard entries. The output shows that the binding entry is configured successfully.

```
[Device] display ipv6 source binding static
Total entries found: 1
MAC Address IP Address VLAN Interface Type
0001-0202-0202 2001::1 N/A GE1/0/1 Static-IPv6
```

## Dynamic IPv6 source guard using DHCPv6 snooping configuration example

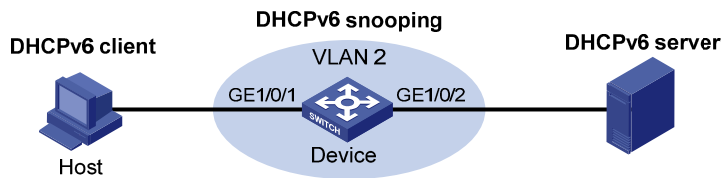
### Network requirements

As shown in [Figure 124](#), the host (DHCPv6 client) and the DHCPv6 server are connected to the device through ports GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 respectively.

Enable DHCPv6 and DHCPv6 snooping on the device, so that the host can obtain an IP address through the DHCPv6 server and the IPv6 IP address and the MAC address of the host can be recorded in a DHCPv6 snooping entry.

Enable IPv6 source guard function on the device's port GigabitEthernet 1/0/1 to filter packets based on DHCPv6 snooping entries, allowing only packets from a client that obtains an IP address through the DHCP server to pass.

**Figure 124 Network diagram**



## Configuration procedure

- Configure DHCPv6 snooping:
  - # Enable DHCPv6 snooping globally.
 

```
<Device> system-view
[Device] ipv6 dhcp snooping enable
```
  - # Enable DHCPv6 snooping in VLAN 2.
 

```
[Device] vlan 2
[Device-vlan2] ipv6 dhcp snooping vlan enable
[Device-vlan2] quit
```
  - # Configure the port connecting to the DHCP server as a trusted port.
 

```
[Device] interface gigabitethernet 1/0/2
[Device-GigabitEthernet1/0/2] ipv6 dhcp snooping trust
[Device-GigabitEthernet1/0/2] quit
```
- Configure the IPv6 source guard function:
  - # Configure the IPv6 source guard function on GigabitEthernet 1/0/1 to filter packets based on both the source IP address and MAC address.
 

```
[Device] interface gigabitethernet 1/0/1
[Device-GigabitEthernet1/0/1] ipv6 verify source ipv6-address mac-address
[Device-GigabitEthernet1/0/1] quit
```

## Verifying the configuration

# Display the dynamic IPv6 source guard entries generated on port GigabitEthernet 1/0/1.

```
[Device] display ipv6 source binding
```

```
Total entries found: 1
```

| MAC Address    | IP Address | VLAN | Interface | Type       |
|----------------|------------|------|-----------|------------|
| 040a-0000-0001 | 2001::1    | 2    | GE1/0/1   | DHCPv6-SNP |

# Display all DHCPv6 snooping entries to see whether they are consistent with the dynamic IP source guard entries generated on GigabitEthernet 1/0/1.

```
[Device] display ipv6 dhcp snooping user-binding dynamic
```

| IP Address | MAC Address    | Lease | VLAN | Interface            |
|------------|----------------|-------|------|----------------------|
| 2001::1    | 040a-0000-0001 | 286   | 2    | GigabitEthernet1/0/1 |

```
--- 1 DHCPv6 snooping item(s) found ---
```

The output shows that a dynamic IPv6 source guard entry has been generated on port GigabitEthernet 1/0/1 based on the DHCPv6 snooping entry.

# Dynamic IPv6 source guard using ND snooping configuration example

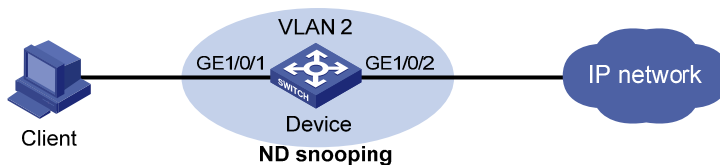
## Network requirements

As shown in Figure 125, the client is connected to the device through port GigabitEthernet 1/0/1.

Enable ND snooping on the device, establishing ND snooping entries by listening to DAD NS messages.

Enable the IPv6 source guard function on port GigabitEthernet 1/0/1 to filter packets based on the ND snooping entries, allowing only packets with a legally obtained IPv6 address to pass.

Figure 125 Network diagram



## Configuration procedure

1. Configure ND snooping:

# In VLAN 2, enable ND snooping.

```
<Device> system-view
[Device] vlan 2
[Device-vlan2] ipv6 nd snooping enable
[Device-vlan2] quit
```

2. Configure the IPv6 source guard function:

# Configure the IPv6 source guard function on GigabitEthernet 1/0/1 to filter packets based on both the source IP address and MAC address.

```
[Device] interface gigabitethernet 1/0/1
[Device-GigabitEthernet1/0/1] ipv6 verify source ipv6-address mac-address
[Device-GigabitEthernet1/0/1] quit
```

## Verifying the configuration

# Display the IPv6 source guard entries generated on port GigabitEthernet 1/0/1.

```
[Device] display ipv6 source binding
```

Total entries found: 1

| MAC Address    | IP Address | VLAN | Interface | Type   |
|----------------|------------|------|-----------|--------|
| 040a-0000-0001 | 2001::1    | 2    | GE1/0/1   | ND-SNP |

# Display the IPv6 ND snooping entries to see whether they are consistent with the dynamic IP source guard entries generated on GigabitEthernet 1/0/1.

```
[Device] display ipv6 nd snooping
```

| IPv6 Address | MAC Address    | VID | Interface | Aging Status |
|--------------|----------------|-----|-----------|--------------|
| 2001::1      | 040a-0000-0001 | 2   | GE1/0/1   | 25 Bound     |

---- Total entries: 1 ----

The output shows that a dynamic IPv6 source guard entry has generated on port GigabitEthernet 1/0/1 based on the ND snooping entry.

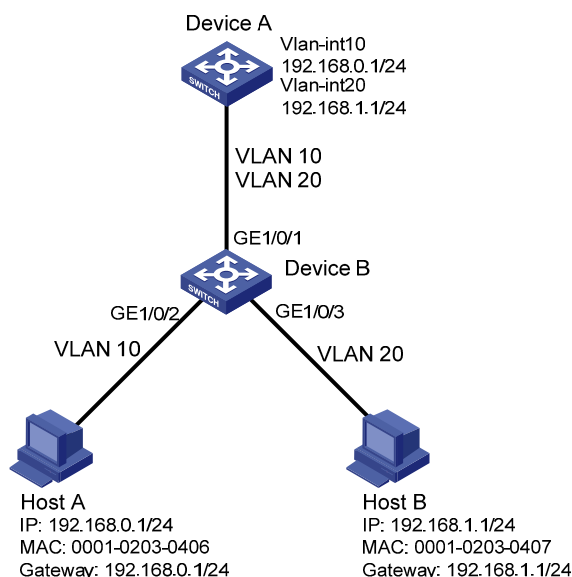
# Global static IP source guard configuration example

## Network requirements

As shown in [Figure 126](#), Device A is a distribution layer device. Device B is an access device. Host A in VLAN 10 and Host B in VLAN 20 communicate with each other through Device A.

- Configure Device B to discard attack packets that exploit the IP address or MAC address of Host A and Host B.
- Configure Device B to forward packets of Host A and Host B normally.

**Figure 126 Network diagram**



## Configuration procedure

# Create VLAN 10, and add port GigabitEthernet 1/0/2 to VLAN 10.

```
<DeviceB> system-view
[DeviceB] vlan 10
[DeviceB-vlan10] port gigabitethernet 1/0/2
[DeviceB-vlan10] quit
```

# Create VLAN 20, and add port GigabitEthernet 1/0/3 to VLAN 20.

```
[DeviceB] vlan 20
[DeviceB-vlan20] port gigabitethernet 1/0/3
[DeviceB-vlan20] quit
```

# Configure the link type of GigabitEthernet 1/0/1 as trunk, and permit packets of VLAN 10 and VLAN 20 to pass the port.

```
[DeviceB] interface gigabitethernet 1/0/1
[DeviceB-GigabitEthernet1/0/1] port link-type trunk
[DeviceB-GigabitEthernet1/0/1] port trunk permit vlan 10 20
[DeviceB-GigabitEthernet1/0/1] quit
```

# Configure IPv4 source guard on GigabitEthernet 1/0/2 and GigabitEthernet 1/0/3 to filter packets based on both the source IP address and MAC address.

```
[DeviceB] interface gigabitethernet 1/0/2
```

```
[DeviceB-GigabitEthernet1/0/2] ip verify source ip-address mac-address
[DeviceB-GigabitEthernet1/0/2] quit
[DeviceB] interface gigabitethernet 1/0/3
[DeviceB-GigabitEthernet1/0/3] ip verify source ip-address mac-address
[DeviceB-GigabitEthernet1/0/3] quit

Configure global static IP binding entries to prevent attack packets that exploit the IP address or MAC
address of Host A and Host B from being forwarded.

[DeviceB] ip source binding ip-address 192.168.0.2 mac-address 0001-0203-0406
[DeviceB] ip source binding ip-address 192.168.1.2 mac-address 0001-0203-0407
```

## Verifying the configuration

# Display static IPv4 binding entries on Device B.

```
[DeviceB] display ip source binding static
```

Total entries found: 2

| MAC Address    | IP Address  | VLAN | Interface | Type   |
|----------------|-------------|------|-----------|--------|
| 0001-0203-0406 | 192.168.0.2 | N/A  | N/A       | Static |
| 0001-0203-0407 | 192.168.1.2 | N/A  | N/A       | Static |

After the configurations, Host A and Host B can ping each other successfully.

# Troubleshooting IP source guard

## Symptom

Failed to configure static or dynamic IP source guard on a port.

## Analysis

IP source guard is not supported on a port in an aggregation group.

## Solution

Remove the port from the aggregation group.

# Configuring ARP attack protection

Only the HP 5500 EI switches support Layer 3 Ethernet port configuration.

The term "interface" in the ARP attack protection features refers to Layer 3 interfaces, including VLAN interfaces and route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).

## Overview

Although ARP is easy to implement, it provides no security mechanism and is vulnerable to network attacks. An attacker can exploit ARP vulnerabilities to attack network devices in the following ways:

- Acts as a trusted user or gateway to send ARP packets so the receiving devices obtain incorrect ARP entries.
- Sends a large number of destination unreachable IP packets to have the receiving device busy with resolving destination IP addresses until its CPU is overloaded.
- Sends a large number of ARP packets to overload the CPU of the receiving device.

For more information about ARP attack features and types, see *ARP Attack Protection Technology White Paper*.

ARP attacks and viruses are threatening LAN security. This chapter introduces multiple features to detect and prevent such attacks.

## ARP attack protection configuration task list

| Task                                 | Remarks                                                                                                                                                    |
|--------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Flood prevention                     | Configuring ARP defense against IP packet attacks<br>Configuring ARP source suppression<br>Optional.<br>Configure this function on gateways (recommended). |
|                                      | Enabling ARP black hole routing<br>Optional.<br>Configure this function on gateways (recommended).                                                         |
|                                      | Configuring ARP packet rate limit<br>Optional.<br>Configure this function on access devices (recommended).                                                 |
|                                      | Configuring source MAC address based ARP attack detection<br>Optional.<br>Configure this function on gateways (recommended).                               |
| User and gateway spoofing prevention | Configuring ARP packet source MAC address consistency check<br>Optional.<br>Configure this function on gateways (recommended).                             |



| Task                                             | Remarks                                                               |
|--------------------------------------------------|-----------------------------------------------------------------------|
| Configuring ARP active acknowledgement           | Optional.<br>Configure this function on gateways (recommended).       |
| Configuring ARP detection                        | Optional.<br>Configure this function on access devices (recommended). |
| Configuring ARP automatic scanning and fixed ARP | Optional.<br>Configure this function on gateways (recommended).       |
| Configuring ARP gateway protection               | Optional.<br>Configure this function on access devices (recommended). |
| Configuring ARP filtering                        | Optional.<br>Configure this function on access devices (recommended). |

## Configuring ARP defense against IP packet attacks

If the device receives a large number of IP packets from a host addressed to unreachable destinations,

- The device sends a large number of ARP requests to the destination subnets, and thus the load of the destination subnets increases.
- The device keeps trying to resolve destination IP addresses, which increases the load on the CPU.

To protect the device from IP packet attacks, you can enable the ARP source suppression function or ARP black hole routing function.

If the packets have the same source address, you can enable the ARP source suppression function. With the function enabled, you can set a threshold for the number of ARP requests that a sending host can trigger in five seconds with packets with unresolvable destination IP addresses. When the number of ARP requests exceeds that threshold, the device suppresses the host from triggering any ARP requests in the following five seconds.

If the packets have various source addresses, you can enable the ARP black hole routing function. After receiving an IP packet whose destination IP address cannot be resolved by ARP, the device with this function enabled immediately creates a black hole route and simply drops all packets matching the route during the aging time of the black hole route.

## Configuring ARP source suppression

| Step                                                                                                                                                                    | Command                                                   | Remarks                     |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------|-----------------------------|
| 1. Enter system view.                                                                                                                                                   | <b>system-view</b>                                        | N/A                         |
| 2. Enable ARP source suppression.                                                                                                                                       | <b>arp source-suppression enable</b>                      | Disabled by default.        |
| 3. Set the maximum number of packets with the same source IP address but unresolvable destination IP addresses that the device can receive in five consecutive seconds. | <b>arp source-suppression limit</b><br><i>limit-value</i> | Optional.<br>10 by default. |

## Enabling ARP black hole routing

| Step                              | Command                           | Remarks                          |
|-----------------------------------|-----------------------------------|----------------------------------|
| 1. Enter system view.             | <b>system-view</b>                | N/A                              |
| 2. Enable ARP black hole routing. | <b>arp resolving-route enable</b> | Optional.<br>Enabled by default. |

## Displaying and maintaining ARP defense against IP packet attacks

| Task                                                          | Command                                                                                                                        | Remarks               |
|---------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Display the ARP source suppression configuration information. | <b>display arp source-suppression</b> [  <br>{ <b>begin</b>   <b>exclude</b>   <b>include</b> }<br><i>regular-expression</i> ] | Available in any view |

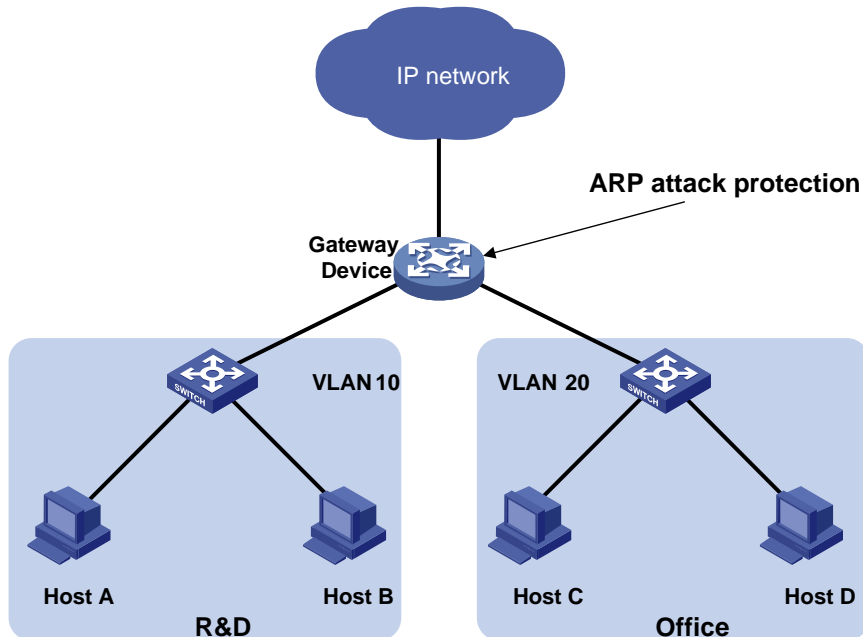
## Configuration example

### Network requirements

As shown in [Figure 127](#), a LAN contains two areas: an R&D area in VLAN 10 and an office area in VLAN 20. The two areas connect to the gateway (Device) through an access switch.

A large number of ARP requests are detected in the office area and are considered as the consequence of an IP flood attack. To prevent such attacks, configure ARP source suppression and ARP black hole routing.

Figure 127 Network diagram



### Configuration considerations

If the attacking packets have the same source address, you can enable the ARP source suppression function with the following steps:

1. Enable ARP source suppression.
2. Set the threshold for ARP packets from the same source address to 100. If the number of ARP requests sourced from the same IP address in five seconds exceeds 100, the device suppresses the IP packets sourced from this IP address from triggering any ARP requests within the following five seconds.

If the attacking packets have different source addresses, enable the ARP black hole routing function on the device.

### Configuration procedure

1. Configure ARP source suppression:

# Enable ARP source suppression on the device and set the threshold for ARP packets from the same source address to 100.

```
<Device> system-view
[Device] arp source-suppression enable
[Device] arp source-suppression limit 100
```

2. Configure ARP black hole routing:

# Enable ARP black hole routing on the device.

```
<Device> system-view
[Device] arp resolving-route enable
```

# Configuring ARP packet rate limit

## Introduction

The ARP packet rate limit feature allows you to limit the rate of ARP packets to be delivered to the CPU on a switch. For example, if an attacker sends a large number of ARP packets to an ARP detection enabled device, the CPU of the device will be overloaded because all of the ARP packets are redirected to the CPU for checking. As a result, the device fails to deliver other functions properly or even crashes. To solve this problem, you can configure ARP packet rate limit.

Enable this feature after the ARP detection or ARP snooping feature is configured, or use this feature to prevent ARP flood attacks.

## Configuration procedure

When the ARP packet rate exceeds the rate limit set on an interface, the device with ARP packet rate limit enabled sends trap and log messages to inform the event. To avoid too many trap and log messages, you can set the interval for sending such messages. Within each interval, the device will output the peak ARP packet rate in the trap and log messages.

Note that trap and log messages are generated only after the trap function of ARP packet rate limit is enabled. Trap and log messages will be sent to the information center of the device. You can set the parameters of the information center to determine the output rules of trap and log messages. The output rules specify whether the messages are allowed to be output and where they are bound for. For the parameter configuration of the information center, see *Network Management and Monitoring Configuration Guide*.

If you enable ARP packet rate limit on a Layer 2 aggregate interface, trap and log messages are sent when the ARP packet rate of a member port exceeds the preset threshold rate.

To configure ARP packet rate limit:

| Step                                                                                                             | Command                                                           | Remarks                                                                                                                                                                     |
|------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                                                                                            | <b>system-view</b>                                                | N/A                                                                                                                                                                         |
| 2. Enable ARP packet rate limit trap.                                                                            | <b>snmp-agent trap enable arp rate-limit</b>                      | Optional.<br>Enabled by default.<br>For more information, see the <b>snmp-agent trap enable arp</b> command in <i>Network Management and Monitoring Command Reference</i> . |
| 3. Set the interval for sending trap and log messages when ARP packet rate exceeds the specified threshold rate. | <b>arp rate-limit information interval seconds</b>                | Optional.<br>60 seconds by default.                                                                                                                                         |
| 4. Enter Layer 2 Ethernet interface/Layer 2 aggregate interface view.                                            | <b>interface</b> <i>interface-type</i><br><i>interface-number</i> | N/A                                                                                                                                                                         |
| 5. Configure ARP packet rate limit.                                                                              | <b>arp rate-limit { disable   rate pps drop }</b>                 | By default, ARP packet rate limit is disabled.                                                                                                                              |

# Configuring source MAC address based ARP attack detection

With this feature enabled, the device checks the source MAC address of ARP packets delivered to the CPU. It detects an attack when one MAC address sends more ARP packets in five seconds than the specified threshold. The device adds the MAC address to the attack detection table.

Before the attack detection entry is aged out, the device uses either of the following detection modes to respond to the detected attack:

- **Monitor mode**—Generates a log message.
- **Filter mode**—Generates a log message and filters out subsequent ARP packets from the attacking MAC address.

You can also configure protected MAC addresses to exclude a gateway or server from detection. A protected MAC address is excluded from ARP attack detection even if it is an attacker.

## Configuration procedure

To configure source MAC address based ARP attack detection:

| Step                                                                                    | Command                                                                    | Remarks                                 |
|-----------------------------------------------------------------------------------------|----------------------------------------------------------------------------|-----------------------------------------|
| 1. Enter system view.                                                                   | <b>system-view</b>                                                         | N/A                                     |
| 2. Enable source MAC address based ARP attack detection and specify the detection mode. | <b>arp anti-attack source-mac { filter   monitor }</b>                     | Disabled by default.                    |
| 3. Configure the threshold.                                                             | <b>arp anti-attack source-mac threshold</b><br><i>threshold-value</i>      | Optional.<br>50 by default.             |
| 4. Configure the age timer for ARP attack detection entries.                            | <b>arp anti-attack source-mac aging-time</b> <i>time</i>                   | Optional.<br>300 seconds by default.    |
| 5. Configure protected MAC addresses.                                                   | <b>arp anti-attack source-mac exclude-mac</b><br><i>mac-address</i> <1-10> | Optional.<br>Not configured by default. |

### NOTE:

After an ARP attack detection entry expires, ARP packets sourced from the MAC address in the entry can be processed normally.

## Displaying and maintaining source MAC address based ARP attack detection

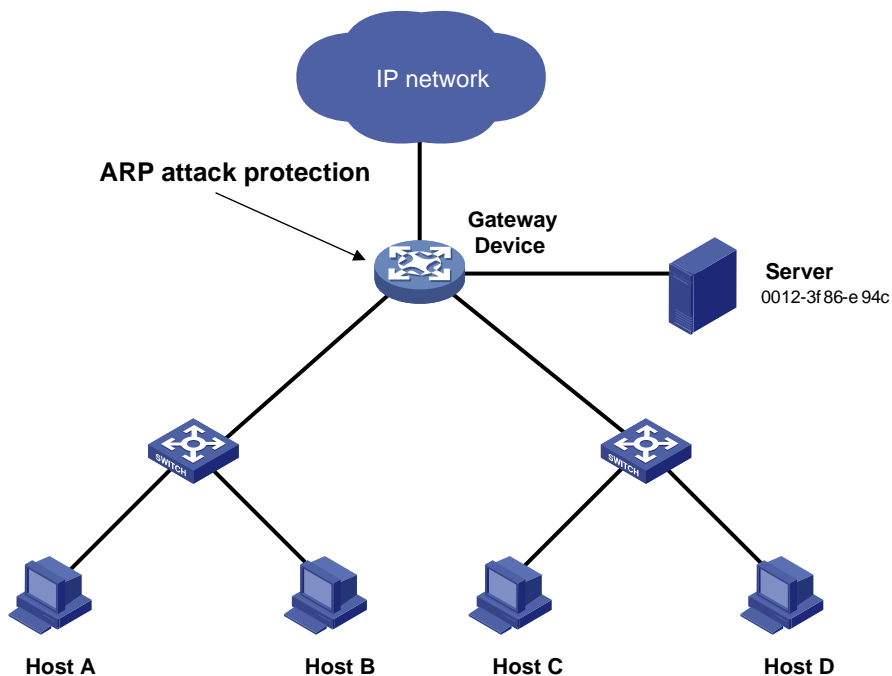
| Task                                                                                       | Command                                                                                                                                                                                       | Remarks               |
|--------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Display attacking MAC addresses detected by source MAC address based ARP attack detection. | <b>display arp anti-attack source-mac</b> { slot slot-number   <b>interface</b> interface-type interface-number } [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } regular-expression ] | Available in any view |

## Configuration example

### Network requirements

As shown in Figure 128, the hosts access the Internet through a gateway (Device). If malicious users send a large number of ARP requests to the gateway, the gateway may crash and cannot process requests from the clients. To solve this problem, configure source MAC address based ARP attack detection on the gateway.

Figure 128 Network diagram



### Configuration considerations

An attacker may forge a large number of ARP packets by using the MAC address of a valid host as the source MAC address. To prevent such attacks, configure the gateway in the following steps:

1. Enable source MAC address based ARP attack detection and specify the filter mode.
2. Set the threshold.
3. Set the age timer for detection entries.
4. Configure the MAC address of the server as a protected MAC address so that it can send ARP packets

### Configuration procedure

# Enable source MAC address based ARP attack detection and specify the filter mode.

```
<Device> system-view
```

```
[Device] arp anti-attack source-mac filter
Set the threshold to 30.
[Device] arp anti-attack source-mac threshold 30
Set the age timer for detection entries to 60 seconds.
[Device] arp anti-attack source-mac aging-time 60
Configure 0012-3f86-e94c as a protected MAC address.
[Device] arp anti-attack source-mac exclude-mac 0012-3f86-e94c
```

## Configuring ARP packet source MAC address consistency check

### Introduction

The ARP packet source MAC address consistency check feature enables a gateway device to filter out ARP packets that have a different source MAC address in the Ethernet header from the sender MAC address in the message, so that the gateway device can learn correct ARP entries.

### Configuration procedure

To enable ARP packet source MAC address consistency check:

| Step                                                       | Command                                   | Remarks             |
|------------------------------------------------------------|-------------------------------------------|---------------------|
| 1. Enter system view.                                      | <b>system-view</b>                        | N/A                 |
| 2. Enable ARP packet source MAC address consistency check. | <b>arp anti-attack valid-check enable</b> | Disabled by default |

## Configuring ARP active acknowledgement

### Introduction

The ARP active acknowledgement feature is configured on gateway devices to identify invalid ARP packets.

ARP active acknowledgement works before the gateway creates or modifies an ARP entry to avoid generating any incorrect ARP entry. For more information about its working mechanism, see *ARP Attack Protection Technology White Paper*.

### Configuration procedure

To configure ARP active acknowledgement:

| Step                  | Command            | Remarks |
|-----------------------|--------------------|---------|
| 1. Enter system view. | <b>system-view</b> | N/A     |

| Step                                               | Command                                  | Remarks             |
|----------------------------------------------------|------------------------------------------|---------------------|
| 2. Enable the ARP active acknowledgement function. | <b>arp anti-attack active-ack enable</b> | Disabled by default |

## Configuring ARP detection

### Introduction

ARP detection enables access devices to block ARP packets from unauthorized clients to prevent user spoofing and gateway spoofing attacks.

ARP detection provides the user validity check, ARP packet validity check, and ARP restricted forwarding functions. If both ARP packet validity check and user validity check are enabled, the former one applies first, and then the latter applies.

ARP detection does not check ARP packets received from ARP trusted ports.

### Configuring user validity check

This feature enables a device to check user validity as follows:

1. Upon receiving an ARP packet from an ARP untrusted interface, the device checks the packet against the configured rules. If a match is found, the ARP packet is processed according to the matching rule; if no match is found, the device checks the packet against static IP Source Guard binding entries
2. The device compares the sender IP and MAC addresses of the ARP packet against the static IP source guard binding entries. If a match is found, the ARP packet is considered valid and is forwarded. If an entry with a matching IP address but an unmatched MAC address is found, the ARP packet is considered invalid and is discarded. If no entry with a matching IP address is found, the device compares the ARP packet's sender IP and MAC addresses against the DHCP snooping entries, 802.1X security entries, and OUI MAC addresses.
3. If a match is found from those entries, the ARP packet is considered valid and is forwarded. (For a packet to pass user validity check based on OUI MAC addresses, the sender MAC address must be an OUI MAC address and the voice VLAN must be enabled.)
4. If no match is found, the ARP packet is considered invalid and is discarded.

For more information about voice VLANs and OUI MAC addresses, see *Layer 2—LAN Switching Configuration Guide*.

#### Configuration guidelines

Follow these guidelines when you configure user validity check:

- Static IP source guard binding entries are created by using the **ip source binding** command. For more information, see "[Configuring IP source guard](#)."
- Dynamic DHCP snooping entries are automatically generated by DHCP snooping. For more information, see *Layer 3—IP Services Configuration Guide*.
- 802.1X security entries are generated by 802.1X. After a client passes 802.1X authentication and uploads its IP address to an ARP detection enabled device, the device automatically generates an 802.1X security entry. Therefore, the 802.1X client must be able to upload its IP address to the device. For more information, see "[Configuring 802.1X](#)."



- At least the configured rules, static IP source guard binding entries, DHCP snooping entries, or 802.1X security entries must be available for user validity check. Otherwise, ARP packets received from ARP untrusted ports will be discarded, except the ARP packets with an OUI MAC address as the sender MAC address when voice VLAN is enabled.
- You must specify a VLAN for an IP source guard binding entry; otherwise, no ARP packets can match the IP source guard binding entry.

## Configuration procedure

To configure user validity check:

| Step                                                                           | Command                                                                                                                                                                                                                                                 | Remarks                                                                                                                                               |
|--------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                                                          | <b>system-view</b>                                                                                                                                                                                                                                      | N/A                                                                                                                                                   |
| 2. Set rules for user validity check.                                          | <b>arp detection</b> <i>id-number</i> { <b>permit</b>   <b>deny</b> } <b>ip</b> { <b>any</b>   <i>ip-address</i> [ <i>ip-address-mask</i> ] } <b>mac</b> { <b>any</b>   <i>mac-address</i> [ <i>mac-address-mask</i> ] } [ <b>vlan</b> <i>vlan-id</i> ] | Optional.<br>By default, no rule is configured.                                                                                                       |
| 3. Enter VLAN view.                                                            | <b>vlan</b> <i>vlan-id</i>                                                                                                                                                                                                                              | N/A                                                                                                                                                   |
| 4. Enable ARP detection for the VLAN.                                          | <b>arp detection enable</b>                                                                                                                                                                                                                             | ARP detection based on static IP source guard binding entries/DHCP snooping entries/802.1X security entries/OUI MAC addresses is disabled by default. |
| 5. Return to system view.                                                      | <b>quit</b>                                                                                                                                                                                                                                             | N/A                                                                                                                                                   |
| 6. Enter Layer 2 Ethernet interface/Layer 2 aggregate interface view.          | <b>interface</b> <i>interface-type</i> <i>interface-number</i>                                                                                                                                                                                          | N/A                                                                                                                                                   |
| 7. Configure the port as a trusted port on which ARP detection does not apply. | <b>arp detection trust</b>                                                                                                                                                                                                                              | Optional.<br>The port is an untrusted port by default.                                                                                                |

## Configuring ARP packet validity check

Perform this task to enable validity check for ARP packets received on untrusted ports and specify the following objects to be checked.

- **src-mac**—Checks whether the sender MAC address in the message body is identical to the source MAC address in the Ethernet header. If they are identical, the packet is forwarded; otherwise, the packet is discarded.
- **dst-mac**—Checks the target MAC address of ARP replies. If the target MAC address is all-zero, all-one, or inconsistent with the destination MAC address in the Ethernet header, the packet is considered invalid and discarded.
- **ip**—Checks the sender and target IP addresses of ARP replies, and the sender IP address of ARP requests. All-zero, all-one, or multicast IP addresses are considered invalid and the corresponding packets are discarded.

To configure ARP packet validity check:

| Step                                                                           | Command                                                                         | Remarks                                                |
|--------------------------------------------------------------------------------|---------------------------------------------------------------------------------|--------------------------------------------------------|
| 1. Enter system view.                                                          | <b>system-view</b>                                                              | N/A                                                    |
| 2. Enter VLAN view.                                                            | <b>vlan</b> <i>vlan-id</i>                                                      | N/A                                                    |
| 3. Enable ARP detection for the VLAN.                                          | <b>arp detection enable</b>                                                     | Disabled by default.                                   |
| 4. Return to system view.                                                      | <b>quit</b>                                                                     | N/A                                                    |
| 5. Enable ARP packet validity check and specify the objects to be checked.     | <b>arp detection validate</b> { <b>dst-mac</b>   <b>ip</b>   <b>src-mac</b> } * | Disabled by default.                                   |
| 6. Enter Layer 2 Ethernet port/Layer 2 aggregate interface view.               | <b>interface</b> <i>interface-type</i><br><i>interface-number</i>               | N/A                                                    |
| 7. Configure the port as a trusted port on which ARP detection does not apply. | <b>arp detection trust</b>                                                      | Optional.<br>The port is an untrusted port by default. |

## Configuring ARP restricted forwarding

ARP restricted forwarding controls the forwarding of ARP packets that are received on untrusted ports and have passed ARP detection in the following cases:

- If the packets are ARP requests, they are forwarded through the trusted ports.
- If the packets are ARP responses, they are forwarded according to their destination MAC address. If no match is found in the MAC address table, they are forwarded through the trusted ports.

Before performing the following configuration, make sure you have configured the **arp detection enable** command.

To enable ARP restricted forwarding:

| Step                                 | Command                                 | Remarks             |
|--------------------------------------|-----------------------------------------|---------------------|
| 1. Enter system view.                | <b>system-view</b>                      | N/A                 |
| 2. Enter VLAN view.                  | <b>vlan</b> <i>vlan-id</i>              | N/A                 |
| 3. Enable ARP restricted forwarding. | <b>arp restricted-forwarding enable</b> | Disabled by default |

## Displaying and maintaining ARP detection

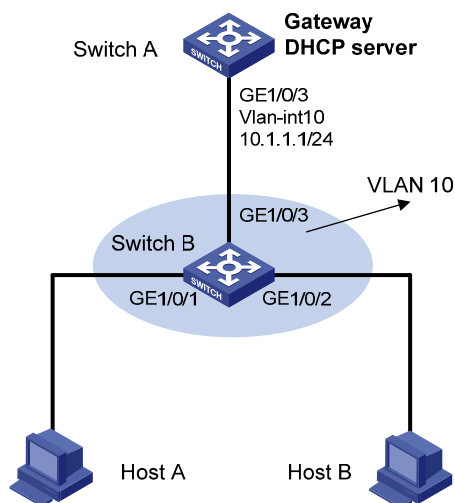
| Task                                          | Command                                                                                                                                                                                         | Remarks                |
|-----------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|
| Display the VLANs enabled with ARP detection. | <b>display arp detection</b> [ [ { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] ]                                                                               | Available in any view  |
| Display the ARP detection statistics.         | <b>display arp detection statistics</b> [ <b>interface</b> <i>interface-type</i> <i>interface-number</i> ] [ [ { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] ] | Available in any view  |
| Clear the ARP detection statistics.           | <b>reset arp detection statistics</b> [ <b>interface</b> <i>interface-type</i> <i>interface-number</i> ]                                                                                        | Available in user view |

# User validity check configuration example

## Network requirements

As shown in Figure 129, configure Switch B to perform user validity check based on 802.1X security entries for connected hosts.

Figure 129 Network diagram



## Configuration procedure

1. Add all the ports on Switch B into VLAN 10, and configure the IP address of VLAN-interface 10 on Switch A. (Details not shown.)

2. Configure Switch A as a DHCP server:

# Configure DHCP address pool 0.

```
<SwitchA> system-view
[SwitchA] dhcp enable
[SwitchA] dhcp server ip-pool 0
[SwitchA-dhcp-pool-0] network 10.1.1.0 mask 255.255.255.0
```

3. Configure Host A and Host B as 802.1X clients and configure them to upload IP addresses for ARP detection. (Details not shown.)

4. Configure Switch B:

# Enable the 802.1X function.

```
<SwitchB> system-view
[SwitchB] dot1x
[SwitchB] interface gigabitEthernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] dot1x
[SwitchB-GigabitEthernet1/0/1] quit
[SwitchB] interface gigabitEthernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] dot1x
[SwitchB-GigabitEthernet1/0/2] quit
```

# Add local access user **test**.

```
[SwitchB] local-user test
[SwitchB-luser-test] service-type lan-access
```

```
[SwitchB-luser-test] password simple test
[SwitchB-luser-test] quit
```

# Enable ARP detection for VLAN 10.

```
[SwitchB] vlan 10
[SwitchB-vlan10] arp detection enable
```

# Configure the upstream port as a trusted port and the downstream ports as untrusted ports (a port is an untrusted port by default).

```
[SwitchB-vlan10] interface gigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] arp detection trust
[SwitchB-GigabitEthernet1/0/3] quit
```

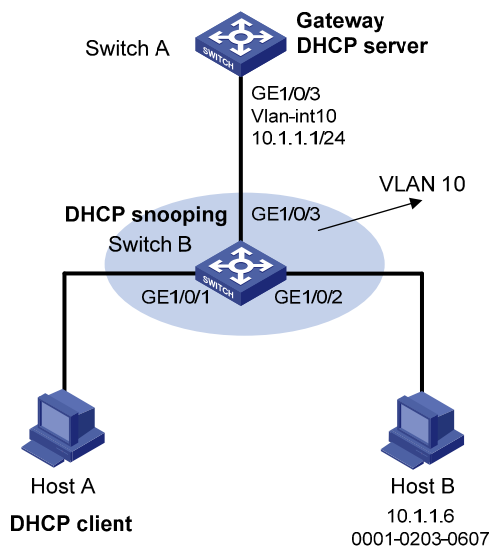
After the preceding configurations are complete, when ARP packets arrive at interfaces GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, they are checked against 802.1X security entries.

## User validity check and ARP packet validity check configuration example

### Network requirements

Configure Switch B to perform ARP packet validity check and user validity check based on static IP source guard binding entries and DHCP snooping entries for connected hosts.

Figure 130 Network diagram



### Configuration procedure

1. Add all the ports on Switch B to VLAN 10, and configure the IP address of VLAN-interface 10 on Switch A. (Details not shown.)
2. Configure Switch A as a DHCP server:

```
Configure DHCP address pool 0.
<SwitchA> system-view
[SwitchA] dhcp enable
[SwitchA] dhcp server ip-pool 0
```

```
[SwitchA-dhcp-pool-0] network 10.1.1.0 mask 255.255.255.0
```

3. Configure Host A as DHCP client, and Host B as user. (Details not shown.)

4. Configure Switch B:

```
Enable DHCP snooping.
```

```
<SwitchB> system-view
```

```
[SwitchB] dhcp-snooping
```

```
[SwitchB] interface gigabitethernet 1/0/3
```

```
[SwitchB-GigabitEthernet1/0/3] dhcp-snooping trust
```

```
[SwitchB-GigabitEthernet1/0/3] quit
```

```
Enable ARP detection for VLAN 10.
```

```
[SwitchB] vlan 10
```

```
[SwitchB-vlan10] arp detection enable
```

```
Configure the upstream port as a trusted port (a port is an untrusted port by default).
```

```
[SwitchB-vlan10] interface gigabitethernet 1/0/3
```

```
[SwitchB-GigabitEthernet1/0/3] arp detection trust
```

```
[SwitchB-GigabitEthernet1/0/3] quit
```

```
Configure a static IP source guard binding entry on interface GigabitEthernet 1/0/2.
```

```
[SwitchB] interface gigabitethernet 1/0/2
```

```
[SwitchB-GigabitEthernet1/0/2] ip source binding ip-address 10.1.1.6 mac-address
0001-0203-0607 vlan 10
```

```
[SwitchB-GigabitEthernet1/0/2] quit
```

```
Enable ARP packet validity check by checking the MAC addresses and IP addresses of ARP
packets.
```

```
[SwitchB] arp detection validate dst-mac ip src-mac
```

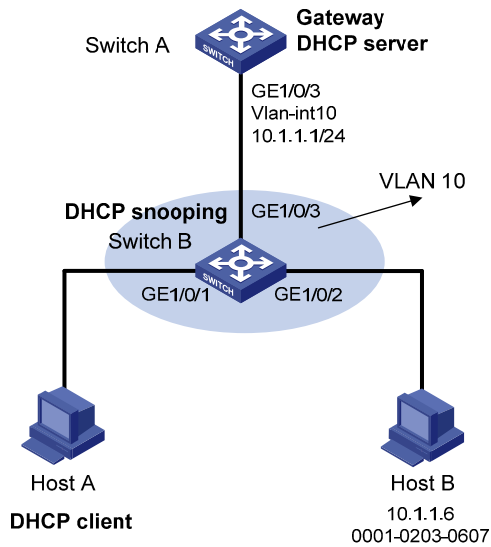
After the configurations are completed, ARP packets received on interfaces GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 have their MAC and IP addresses checked first, and then are checked against the static IP source guard binding entries and finally DHCP snooping entries.

## ARP restricted forwarding configuration example

### Network requirements

As shown in [Figure 131](#), configure ARP restricted forwarding on Switch B where ARP detection is configured so that port isolation configured on Switch B can take effect for broadcast ARP requests.

**Figure 131 Network diagram**



### Configuration procedure

1. Configure VLAN 10, add ports to VLAN 10, and configure the IP address of the VLAN-interface, as shown in [Figure 127](#). (Details not shown.)
2. Configure the DHCP server on Switch A.

# Configure DHCP address pool 0.

```
<SwitchA> system-view
[SwitchA] dhcp enable
[SwitchA] dhcp server ip-pool 0
[SwitchA-dhcp-pool-0] network 10.1.1.0 mask 255.255.255.0
```

3. Configure the DHCP client on Hosts A and B. (Details not shown.)

4. Configure Switch B.

# Enable DHCP snooping, and configure GigabitEthernet 1/0/3 as a DHCP-trusted port.

```
<SwitchB> system-view
[SwitchB] dhcp-snooping
[SwitchB] interface gigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] dhcp-snooping trust
[SwitchB-GigabitEthernet1/0/3] quit
```

# Enable ARP detection.

```
[SwitchB] vlan 10
[SwitchB-vlan10] arp detection enable
```

# Configure GigabitEthernet 1/0/3 as an ARP-trusted port.

```
[SwitchB-vlan10] interface gigabitEthernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] arp detection trust
[SwitchB-GigabitEthernet1/0/3] quit
```

# Configure a static IP source guard entry on interface GigabitEthernet 1/0/2.

```
[SwitchB] interface gigabitEthernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] ip source binding ip-address 10.1.1.6 mac-address
0001-0203-0607 vlan 10
[SwitchB-GigabitEthernet1/0/2] quit
```

# Enable the checking of the MAC addresses and IP addresses of ARP packets.

```
[SwitchB] arp detection validate dst-mac ip src-mac
```

# Configure port isolation.

```
[SwitchB] interface gigabitethernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] port-isolate enable
[SwitchB-GigabitEthernet1/0/1] quit
[SwitchB] interface gigabitethernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] port-isolate enable
[SwitchB-GigabitEthernet1/0/2] quit
```

After the preceding configurations are complete, ARP packets received on interfaces GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 have their MAC and IP addresses checked first, and then are checked against the static IP source guard binding entries and finally DHCP snooping entries. However, ARP broadcast requests sent from Host A can pass the check on Switch B and reach Host B. Port isolation fails.

# Configure ARP restricted forwarding.

```
[SwitchB] vlan 10
[SwitchB-vlan10] arp restricted-forwarding enable
[SwitchB-vlan10] quit
```

After the configuration, Switch B forwards ARP broadcast requests from Host A to Switch A through the trusted port GigabitEthernet 1/0/3, and thus Host B cannot receive such packets. Port isolation works normally.

## Configuring ARP automatic scanning and fixed ARP

ARP automatic scanning is usually used together with the fixed ARP feature.

With ARP automatic scanning enabled on an interface, the device automatically scans neighbors on the interface, sends ARP requests to the neighbors, obtains their MAC addresses, and creates dynamic ARP entries.

Fixed ARP allows the device to change the existing dynamic ARP entries (including those generated through ARP automatic scanning) into static ARP entries. The fixed ARP feature effectively prevents ARP entries from being modified by attackers.

HP recommends that you use ARP automatic scanning and fixed ARP in a small-scale network such as a cybercafe.

## Configuration guidelines

Follow these guidelines when you configure ARP automatic scanning and fixed ARP:

- IP addresses existing in ARP entries are not scanned.
- ARP automatic scanning may take some time. To stop an ongoing scan, press **Ctrl + C**. Dynamic ARP entries are created based on ARP replies received before the scan is terminated.
- The static ARP entries changed from dynamic ARP entries have the same attributes as the manually configured static ARP entries.
- Use the **arp fixup** command to change the existing dynamic ARP entries into static ARP entries. You can use this command again to change the dynamic ARP entries learned later into static ARP entries.

- The number of static ARP entries changed from dynamic ARP entries is restricted by the number of static ARP entries that the device supports. As a result, the device may fail to change all dynamic ARP entries into static ARP entries.
- To delete a specific static ARP entry changed from a dynamic one, use the **undo arp ip-address** [ *vpn-instance-name* ] command (The HP 5500 SI switch series does not support the *vpn-instance-name* argument in the command). To delete all such static ARP entries, use the **reset arp all** or **reset arp static** command.

## Configuration procedure

To configure ARP automatic scanning and fixed ARP:

| Step                              | Command                                                                     |
|-----------------------------------|-----------------------------------------------------------------------------|
| 1. Enter system view.             | <b>system-view</b>                                                          |
| 2. Enter interface view.          | <b>interface</b> <i>interface-type</i> <i>interface-number</i>              |
| 3. Enable ARP automatic scanning. | <b>arp scan</b> [ <i>start-ip-address</i> <b>to</b> <i>end-ip-address</i> ] |
| 4. Return to system view.         | <b>quit</b>                                                                 |
| 5. Enable fixed ARP.              | <b>arp fixup</b>                                                            |

## Configuring ARP gateway protection

The ARP gateway protection feature, if configured on ports not connected with the gateway, can block gateway spoofing attacks.

When such a port receives an ARP packet, it checks whether the sender IP address in the packet is consistent with that of any protected gateway. If yes, it discards the packet. If not, it handles the packet normally.

## Configuration guidelines

Follow these guidelines when you configure ARP gateway protection:

- You can enable ARP gateway protection for up to eight gateways on a port.
- Commands **arp filter source** and **arp filter binding** cannot be both configured on a port.
- If ARP gateway protection works with ARP detection and ARP snooping, ARP gateway protection applies first.

## Configuration procedure

To configure ARP gateway protection:

| Step                                                                       | Command                                                        | Remarks |
|----------------------------------------------------------------------------|----------------------------------------------------------------|---------|
| 1. Enter system view.                                                      | <b>system-view</b>                                             | N/A     |
| 2. Enter Layer 2 Ethernet interface view/Layer 2 aggregate interface view. | <b>interface</b> <i>interface-type</i> <i>interface-number</i> | N/A     |



| Step | Command                                                                                                 | Remarks             |
|------|---------------------------------------------------------------------------------------------------------|---------------------|
| 3.   | Enable ARP gateway protection for a specified gateway.<br><br><code>arp filter source ip-address</code> | Disabled by default |

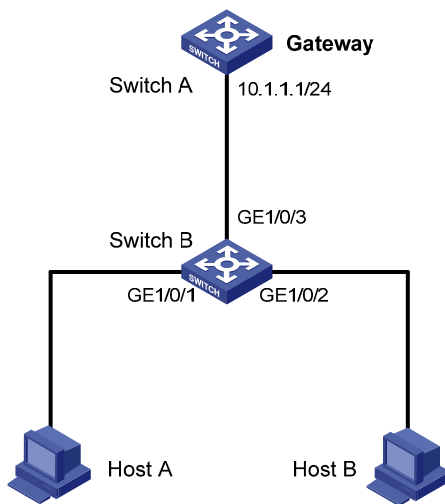
## Configuration example

### Network requirements

As shown in [Figure 132](#), Host B launches gateway spoofing attacks to Switch B. As a result, traffic that Switch B intends to send to Switch A is sent to Host B.

Configure Switch B to block such attacks.

**Figure 132 Network diagram**



### Configuration procedure

# Configure ARP gateway protection on Switch B.

```

<SwitchB> system-view
[SwitchB] interface gigabitethernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] arp filter source 10.1.1.1
[SwitchB-GigabitEthernet1/0/1] quit
[SwitchB] interface gigabitethernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] arp filter source 10.1.1.1

```

After the configuration is complete, Switch B will discard the ARP packets whose source IP address is that of the gateway.

## Configuring ARP filtering

To prevent gateway spoofing and user spoofing, the ARP filtering feature controls the forwarding of ARP packets on a port.

The port checks the sender IP and MAC addresses in a received ARP packet against configured ARP filtering entries. If a match is found, the packet is handled normally. If not, the packet is discarded.

## Configuration guidelines

Follow these guidelines when you configure ARP filtering:

- You can configure up to eight ARP filtering entries on a port.
- Commands **arp filter source** and **arp filter binding** cannot be both configured on a port.
- If ARP filtering works with ARP detection and ARP snooping, ARP filtering applies first.

## Configuration procedure

To configure ARP filtering:

| Step                                                                       | Command                                                           | Remarks                   |
|----------------------------------------------------------------------------|-------------------------------------------------------------------|---------------------------|
| 1. Enter system view.                                                      | <b>system-view</b>                                                | N/A                       |
| 2. Enter Layer 2 Ethernet interface view/Layer 2 aggregate interface view. | <b>interface</b> <i>interface-type</i><br><i>interface-number</i> | N/A                       |
| 3. Configure an ARP filtering entry.                                       | <b>arp filter binding</b> <i>ip-address</i><br><i>mac-address</i> | Not configured by default |

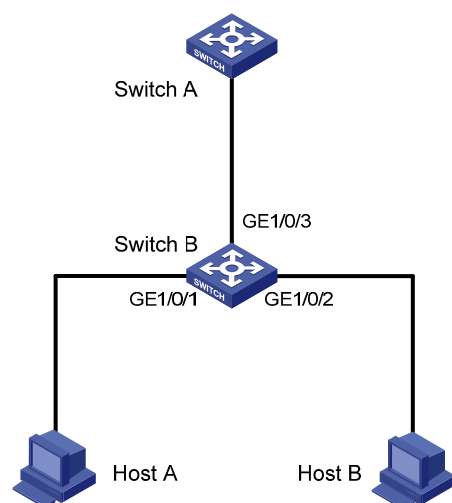
## Configuration example

### Network requirements

As shown in [Figure 133](#), the IP and MAC addresses of Host A are 10.1.1.2 and 000f-e349-1233. The IP and MAC addresses of Host B are 10.1.1.3 and 000f-e349-1234.

Configure ARP filtering on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 of Switch B to permit specific ARP packets only.

**Figure 133 Network diagram**



### Configuration procedure

# Configure ARP filtering on Switch B.

```
<SwitchB> system-view
```

```
[SwitchB] interface gigabitethernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] arp filter binding 10.1.1.2 000f-e349-1233
[SwitchB-GigabitEthernet1/0/1] quit
[SwitchB] interface gigabitethernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] arp filter binding 10.1.1.3 000f-e349-1234
```

After the configuration is complete, GigabitEthernet 1/0/1 will permit incoming ARP packets with sender IP and MAC addresses as 10.1.1.2 and 000f-e349-1233, and discard other ARP packets. GigabitEthernet 1/0/2 will permit incoming ARP packets with sender IP and MAC addresses as 10.1.1.9 and 000f-e349-1233 and discard other ARP packets. ARP packets from Host A are permitted, but those from Host B are discarded.

# Configuring ND attack defense

## Overview

The IPv6 Neighbor Discovery (ND) protocol provides rich functions, such as address resolution, neighbor reachability detection, duplicate address detection, router/prefix discovery and address autoconfiguration, and redirection. However, it does not provide any security mechanisms. Attackers can easily exploit the ND protocol to attack hosts and gateways by sending forged packets. For more information about the five functions of the ND protocol, see *Layer 3—IP Services Configuration Guide*.

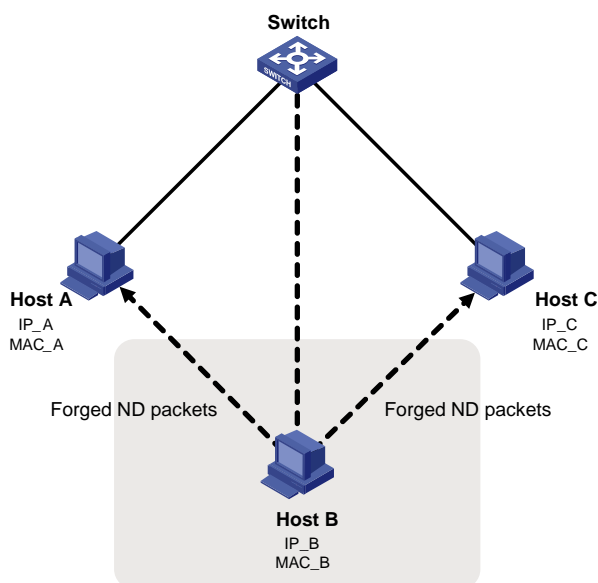
The ND protocol implements its function by using five types of ICMPv6 messages:

- Neighbor Solicitation (NS)
- Neighbor Advertisement (NA)
- Router Solicitation (RS)
- Router Advertisement (RA)
- Redirect (RR)

As shown in [Figure 134](#), an attacker can attack a network by sending forged ICMPv6 messages:

- Sends forged NS/NA/RS packets with the IPv6 address of a victim host. The gateway and other hosts update the ND entry for the victim host with incorrect address information. As a result, all packets intended for the victim host are sent to the attacking host rather than the victim host.
- Sends forged RA packets with the IPv6 address of a victim gateway. As a result, all hosts attached to the victim gateway maintain incorrect IPv6 configuration parameters and ND entries.

**Figure 134 ND attack diagram**



All forged ND packets have two common features:

- The Ethernet frame header and the source link layer address option of the ND packet contain different source MAC addresses.

- The mapping between the source IPv6 address and the source MAC address in the Ethernet frame header is invalid.

To identify forged ND packets, HP developed the source MAC consistency check and ND detection features.

## Enabling source MAC consistency check for ND packets

Use source MAC consistency check on a gateway to filter out ND packets that carry different source MAC addresses in the Ethernet frame header and the source link layer address option.

Follow these guidelines when you enable source MAC consistency check for ND packets:

- If VRRP is used, disable source MAC consistency check for ND packets to prevent incorrect dropping of packets. With VRRP, the NA message always conveys a MAC address different from the Source Link-Layer Address option.
- Only the HP 5500 EI switches support VRRP.

To enable source MAC consistency check for ND packets:

| Step                                                   | Command                         | Remarks             |
|--------------------------------------------------------|---------------------------------|---------------------|
| 1. Enter system view.                                  | <b>system-view</b>              | N/A                 |
| 2. Enable source MAC consistency check for ND packets. | <b>ipv6 nd mac-check enable</b> | Disabled by default |

## Configuring the ND detection function

### Introduction to ND detection

Use the ND detection function on access devices to verify the source of ND packets. If an ND packet comes from a spoofing host or gateway, it is discarded.

The ND detection function operates on a per VLAN basis. In an ND detection-enabled VLAN, a port is either ND-trusted or ND-untrusted:

- An ND-trusted port does not check ND packets for address spoofing.
- An ND-untrusted port checks all ND packets but RA and RR messages in the VLAN for source spoofing. RA and RR messages are considered illegal and are discarded directly.

The ND detection function checks an ND packet by looking up the IPv6 static bindings table of the IP source guard function, ND snooping table, and DHCPv6 snooping table in the following steps:

1. Looks up the IPv6 static binding table of IP source guard, based on the source IPv6 address and the source MAC address in the Ethernet frame header of the ND packet. If an exact match is found, the ND packet is forwarded. If an entry matches the source IPv6 address but not the source MAC address, the ND packet is discarded. If no entry matches the source IPv6 address, the ND detection function continues to look up the DHCPv6 snooping table and the ND snooping table.
2. If an exact match is found in either the DHCPv6 snooping or ND snooping table, the ND packet is forwarded. If no match is found in either table, the packet is discarded. If neither the DHCPv6 snooping table nor the ND snooping table is available, the ND packet is discarded.

## Configuration guidelines

Follow these guidelines when you configure ND detection:

- To create IPv6 static bindings with IP source guard, use the **ipv6 source binding** command. For more information, see "[Configuring IP source guard](#)."
- The DHCPv6 snooping table is created automatically by the DHCPv6 snooping module. For more information, see *Layer 3—IP Services Configuration Guide*.
- The ND snooping table is created automatically by the ND snooping module. For more information, see *Layer 3—IP Services Configuration Guide*.
- ND detection performs source check by using the binding tables of IP source guard, DHCPv6 snooping, and ND snooping. To prevent an ND-untrusted port from discarding legal ND packets in an ND detection-enabled VLAN, make sure that at least one of the three functions is available.
- When creating an IPv6 static binding with IP source guard for ND detection in a VLAN, specify the VLAN ID for the binding. If not, no ND packets in the VLAN can match the binding.

## Configuration procedure

To configure ND detection:

| Step                                                                          | Command                                                           | Remarks                                                              |
|-------------------------------------------------------------------------------|-------------------------------------------------------------------|----------------------------------------------------------------------|
| 1. Enter system view.                                                         | <b>system-view</b>                                                | N/A                                                                  |
| 2. Enter VLAN view.                                                           | <b>vlan</b> <i>vlan-id</i>                                        | N/A                                                                  |
| 3. Enable ND Detection.                                                       | <b>ipv6 nd detection enable</b>                                   | Disabled by default.                                                 |
| 4. Quit system view.                                                          | <b>quit</b>                                                       | N/A                                                                  |
| 5. Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view. | <b>interface</b> <i>interface-type</i><br><i>interface-number</i> | N/A                                                                  |
| 6. Configure the port as an ND-trusted port.                                  | <b>ipv6 nd detection trust</b>                                    | Optional.<br>A port does not trust sources of ND packets by default. |

## Displaying and maintaining ND detection

| Task                                                                                        | Command                                                                                                                                                                                           | Remarks                |
|---------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|
| Display the ND detection configuration.                                                     | <b>display ipv6 nd detection</b> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                                                                               | Available in any view  |
| Display the statistics of discarded packets when the ND detection checks the user legality. | <b>display ipv6 nd detection statistics</b> [ <b>interface</b> <i>interface-type</i> <i>interface-number</i> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] | Available in any view  |
| Clear the statistics by ND detection.                                                       | <b>reset ipv6 nd detection statistics</b> [ <b>interface</b> <i>interface-type</i> <i>interface-number</i> ]                                                                                      | Available in user view |

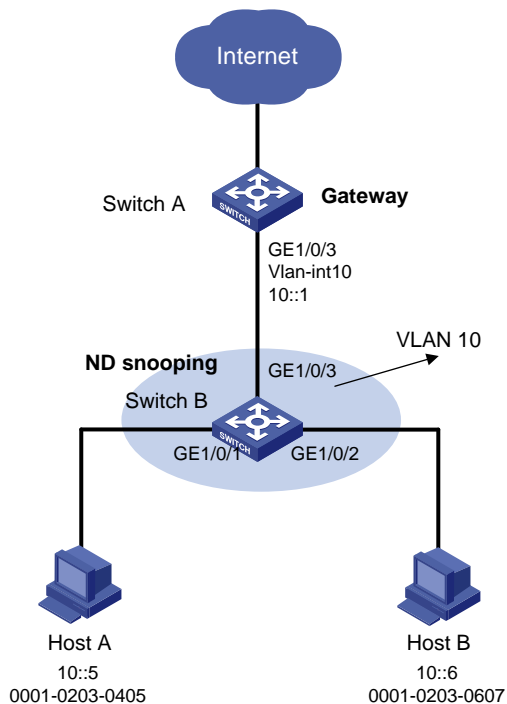
# ND detection configuration example

## Network requirements

As shown in [Figure 135](#), Host A and Host B connect to Switch A, the gateway, through Switch B. Host A has the IPv6 address 10::5 and MAC address 0001-0203-0405. Host B has the IPv6 address 10::6 and MAC address 0001-0203-0607.

Enable ND detection on Switch B to filter out forged ND packets.

**Figure 135 Network diagram**



## Configuration procedure

### 1. Configuring Switch A:

# Enable IPv6 forwarding.

```
<SwitchA> system-view
```

```
[SwitchA] ipv6
```

# Create VLAN 10.

```
[SwitchA] vlan 10
```

```
[SwitchA-vlan10] quit
```

# Assign port GigabitEthernet 1/0/3 to VLAN 10.

```
[SwitchA] interface gigabitethernet 1/0/3
```

```
[SwitchA- GigabitEthernet1/0/3] port link-type trunk
```

```
[SwitchA- GigabitEthernet1/0/3] port trunk permit vlan 10
```

```
[SwitchA- GigabitEthernet1/0/3] quit
```

# Assign an IPv6 address to VLAN-interface 10.

```
[SwitchA] interface vlan-interface 10
```

```
[SwitchA-Vlan-interface10] ipv6 address 10::1/64
[SwitchA-Vlan-interface10] quit
```

## 2. Configuring Switch B:

**# Enable IPv6 forwarding.**

```
<SwitchB> system-view
[SwitchB] ipv6
```

**# Create VLAN 10.**

```
[SwitchB] vlan 10
[SwitchB-vlan10] quit
```

**# Add ports GigabitEthernet 1/0/1 through GigabitEthernet 1/0/3 to VLAN 10.**

```
[SwitchB] interface gigabitethernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] port access vlan 10
[SwitchB-GigabitEthernet1/0/1] quit
[SwitchB] interface gigabitethernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] port access vlan 10
[SwitchB-GigabitEthernet1/0/2] quit
[SwitchB] interface gigabitethernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] port link-type trunk
[SwitchB-GigabitEthernet1/0/3] port trunk permit vlan 10
[SwitchB-GigabitEthernet1/0/3] quit
```

**# Enable ND snooping for global unicast and link local addresses in VLAN 10.**

```
[SwitchB] ipv6 nd snooping enable link-local
[SwitchB] ipv6 nd snooping enable global
[SwitchB] vlan 10
[SwitchB-vlan 10] ipv6 nd snooping enable
```

**# Enable ND detection in VLAN 10.**

```
[SwitchB-vlan 10] ipv6 nd detection enable
[SwitchB-vlan 10] quit
```

**# Configure the uplink port GigabitEthernet 1/0/3 as an ND-trusted port, and the downlink ports GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 as ND-untrusted ports (the default).**

```
[SwitchB] interface gigabitethernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] ipv6 nd detection trust
```

The configuration enables Switch B to check all incoming ND packets of ports GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 based on the ND snooping table.



# Configuring URPF (available only on the HP 5500 EI)

The term "router" in this feature refers to both routers and Layer 3 switches.

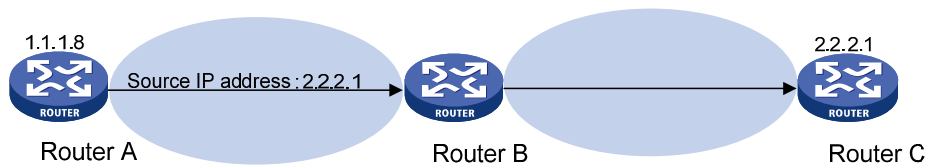
## URPF overview

### What is URPF

Unicast Reverse Path Forwarding (URPF) protects a network against source address spoofing attacks, such as denial of service (DoS) and distributed denial of service (DDoS) attacks.

Attackers launch attacks by creating a series of packets with forged source addresses. For applications using IP-address-based authentication, this type of attack allows unauthorized users to access the system in the name of authorized users, or to even access the system as the administrator. Even if the attackers cannot receive any response packets, the attacks are still disruptive to the attacked target.

**Figure 136 Attack based on source address spoofing**



As shown in [Figure 136](#), Router A sends the server (Router B) requests with a forged source IP address 2.2.2.1 at a high rate, and Router B sends packets to IP address 2.2.2.1 (Router C) in response to the requests. Consequently, both Router B and Router C are attacked.

URPF can prevent this source address spoofing attack by checking the source addresses of packets and filtering out invalid packets.

## URPF check modes

URPF provides two check modes: strict and loose.

### Strict URPF

To pass strict URPF check, the source address and receiving interface of a packet must match the destination address and output interface of a forwarding information base (FIB) entry.

In some scenarios such as asymmetrical routing, strict URPF may discard valid packets.

Strict URPF is often deployed between an internet service provider (ISP) and the connected users.

### Loose URPF

To pass loose URPF check, the source address of a packet must match the destination address of a FIB entry. Loose URPF can avoid discarding valid packets, but may let go attack packets.

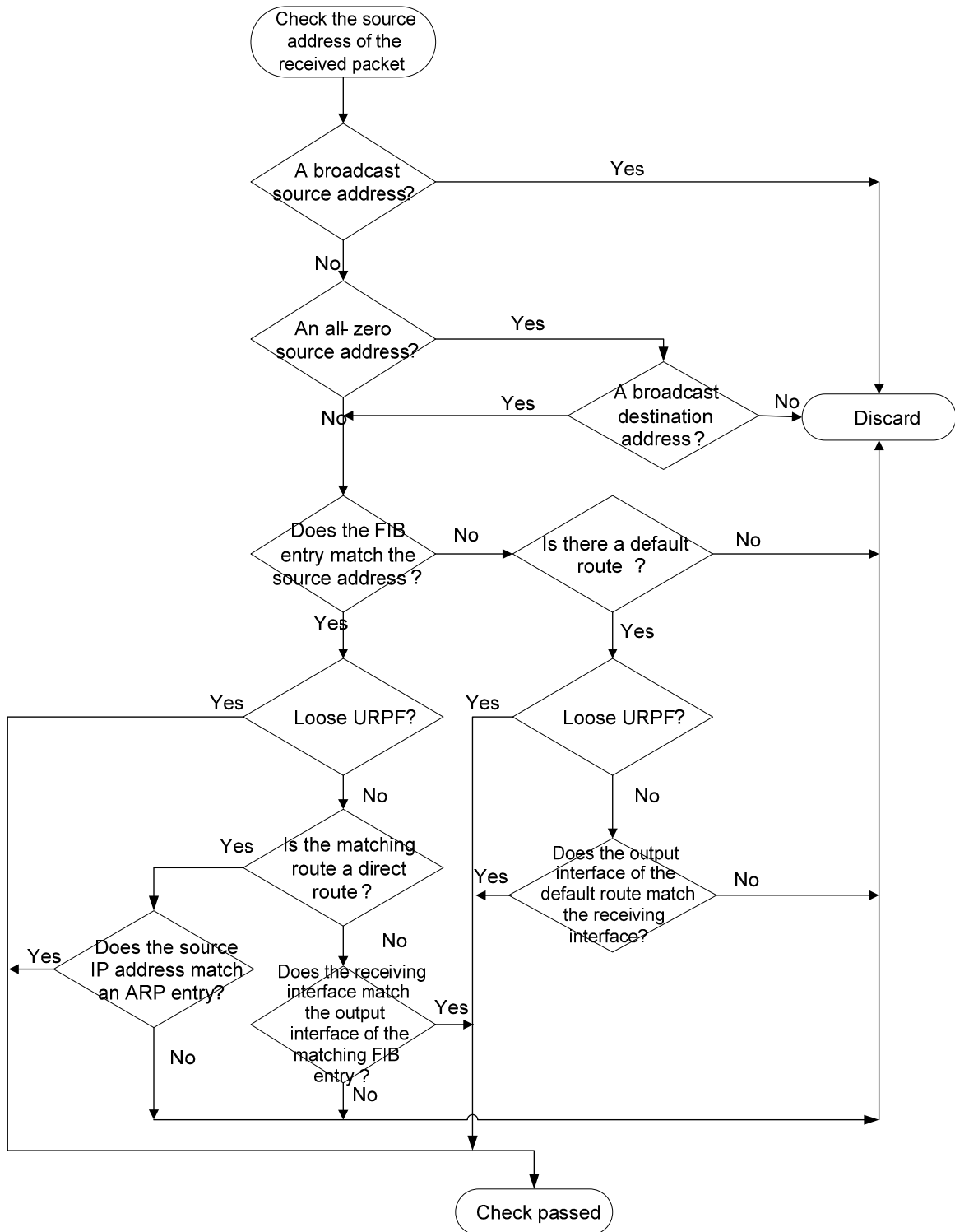
Loose URPF is often deployed between ISPs, especially in asymmetrical routing.

# How URPF works

URPF does not check multicast packets.

URPF works in the steps, as shown in [Figure 137](#).

**Figure 137 URPF work flow**

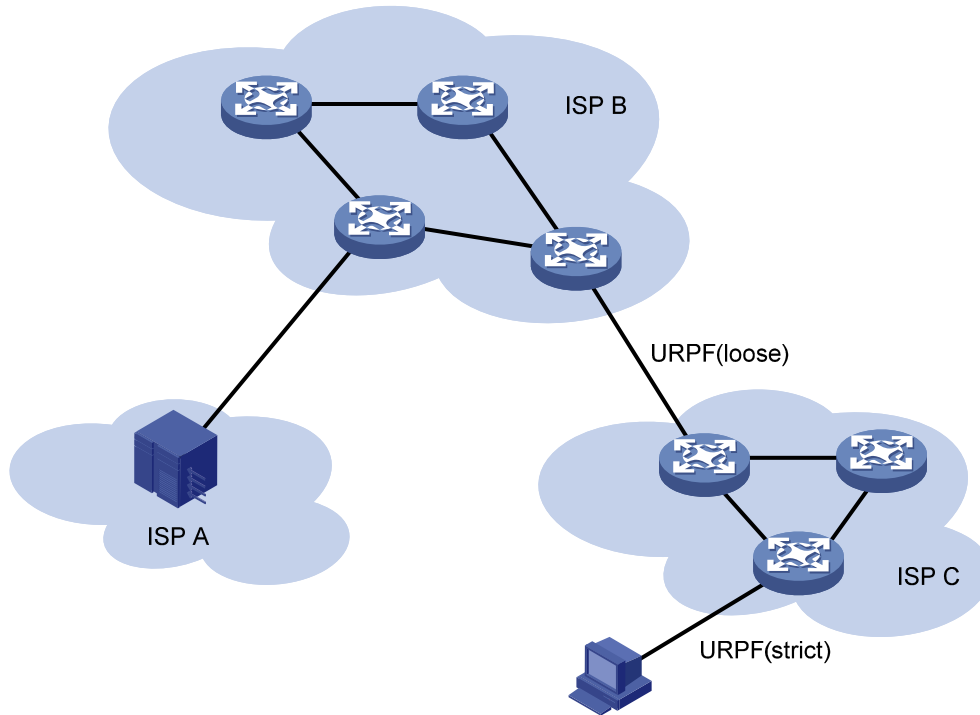


1. URPF checks the source address validity:

- Discards packets with a broadcast source address.
  - Discards packets with an all-zero source address but a non-broadcast destination address. (A packet with source address 0.0.0.0 and destination address 255.255.255.255 might be a DHCP or BOOTP packet, and is not discarded.)
  - For other packets, precede to step 2.
2. URPF checks whether the source address matches a FIB entry:
    - If yes, precede to step 3.
    - If not, precede to step 6.
  3. URPF checks whether the check mode is loose:
    - If yes, precede to step 8.
    - If not, URPF checks whether the matching route is a direct route: if yes, precede to step 5; if not, precede to step 4.
  4. URPF checks whether the receiving interface matches the output interface of the matching FIB entry.
    - If yes, precede to step 8.
    - If not, precede to step 9.
  5. URPF checks whether the source IP address matches an ARP entry.
    - If yes, precede to step 8.
    - If not, precede to step 9.
  6. URPF checks whether the FIB table has a default route.
    - If yes, precede to step 7.
    - If not, precede to step 9.
  7. URPF checks whether the check mode is loose:
    - If yes, precede to step 8.
    - If not, URPF checks whether the output interface of the default route matches the receiving interface of the packet: if yes, precede to step 8; if not, precede to step 9.
  8. The packet passes the check and is forwarded.
  9. The packet is discarded.

# Network application

Figure 138 Network diagram



Configure strict URPF between each ISP and its connected users, and loose URPF between ISPs.

## Configuring URPF

To configure URPF globally:

| Step                           | Command                           | Remarks              |
|--------------------------------|-----------------------------------|----------------------|
| 1. Enter system view.          | <b>system-view</b>                | N/A                  |
| 2. Enable URPF check globally. | <b>ip urpf { loose   strict }</b> | Disabled by default. |

### NOTE:

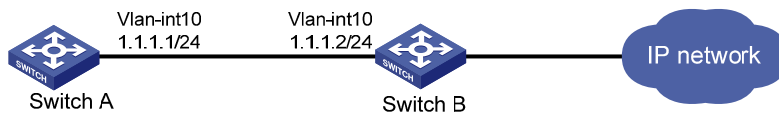
- The routing table size decreases by half when URPF is enabled on the HP 5500 EI switches.
- To prevent loss of routes and packets, URPF cannot be enabled if the number of route entries the switch maintains exceeds half the routing table size.

## URPF configuration example

### Network requirements

As shown in Figure 139, a client (Switch A) directly connects to the ISP switch (Switch B). Enable URPF check on Switch A and Switch B to prevent source address spoofing attacks.

**Figure 139 Network diagram**



### Configuration procedure

1. Configure Switch A:  
# Enable strict URPF check.  
<SwitchA> system-view  
[SwitchA] ip urpf strict
2. Configure Switch B:  
# Enable strict URPF check.  
<SwitchB> system-view  
[SwitchB] ip urpf strict

# Configuring SAVI

## SAVI overview

Source Address Validation (SAVI) is applied on access devices. SAVI creates a table of bindings between addresses and ports through other features such as ND snooping, DHCPv6 snooping, and IP Source Guard, and uses those bindings to check the validity of the source addresses of DHCPv6 protocol packets, ND protocol packets, and IPv6 data packets.

SAVI can be used in the following address assignment scenarios:

- DHCPv6-only: The hosts connected to the SAVI-enabled device obtain addresses only through DHCPv6.
- SLAAC-only: The hosts connected to the SAVI-enabled device obtain addresses only through Stateless Address Autoconfiguration (SLAAC).
- DHCPv6+SLAAC: The hosts connected to the SAVI-enabled device obtain addresses through DHCPv6 and SLAAC.

The following section describes SAVI configurations in these address assignment scenarios.

## Configuring global SAVI

| Step                                                                | Command                                        | Remarks                                                                                                                                                                                                                                                                                                                                                                 |
|---------------------------------------------------------------------|------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                                               | <b>system-view</b>                             | N/A                                                                                                                                                                                                                                                                                                                                                                     |
| 2. Enable the SAVI function.                                        | <b>ipv6 savi strict</b>                        | Disabled by default.                                                                                                                                                                                                                                                                                                                                                    |
| 3. Set the time to wait for a duplicate address detection (DAD) NA. | <b>ipv6 savi dad-delay</b> <i>value</i>        | Optional<br>One second by default.<br>If no DAD NA is received within the specified time when the corresponding ND snooping entry is in detect state, the ND snooping entry changes to bound state.                                                                                                                                                                     |
| 4. Set the time to wait for a DAD NS from a DHCPv6 client.          | <b>ipv6 savi dad-preparedelay</b> <i>value</i> | Optional<br>One second by default.<br>This command is used with the DHCPv6 snooping function. After DHCPv6 snooping detects that a client obtains an IPv6 address, it monitors whether the client detects IP address conflict. If DHCPv6 snooping does not receive any DAD NS from the client before the set time expires, SAVI sends a DAD NS on behalf of the client. |

---

**NOTE:**

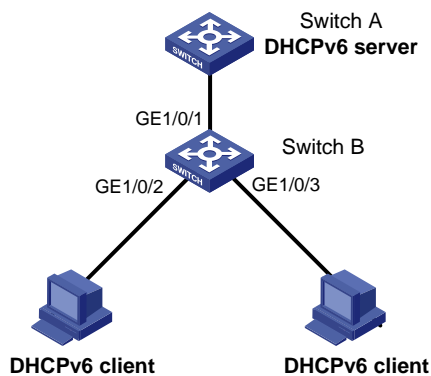
If a port on the SAVI enabled device is down for three minutes or more, the device deletes the DHCPv6 snooping entries and ND snooping entries corresponding to the port.

---

# SAVI configuration in DHCPv6-only address assignment scenario

## Network requirements

Figure 140 Network diagram



As shown in Figure 140, Switch A is the DHCPv6 server. Switch B connects to the DHCPv6 server through interface GigabitEthernet 1/0/1, and connects to two DHCPv6 clients through interfaces GigabitEthernet 1/0/2 and GigabitEthernet 1/0/3. The three interfaces of Switch B belong to VLAN 2. The client can obtain IP address only through DHCPv6. Configure SAVI on Switch B to automatically bind the IP addresses assigned through DHCPv6 and permit only packets from bound addresses and link-local addresses.

## Configuration considerations

Configure Switch B as follows:

- Enable SAVI.
- Enable DHCPv6 snooping. For more information about DHCPv6 snooping, see *Layer 3—IP Services Configuration Guide*.
- Enable link-local address ND snooping. For more information about ND snooping, see *Layer 3—IP Services Configuration Guide*.
- Enable ND detection in VLAN 2 to check the ND packets arrived on the ports. For more information about ND detection, see "[Configuring ND attack defense](#)."
- Configure a static IPv6 source guard binding entry on each interface connected to a client. This step is optional. If this step is not performed, SAVI does not check packets against static binding entries. For more information about static IPv6 source guard binding entries, see "[Configuring IP source guard](#)."
- Configure dynamic IPv6 source guard binding on the interfaces connected to the clients. For more information about dynamic IPv6 source guard binding, see "[Configuring IP source guard](#)."

## Packet check principles

Switch B checks DHCPv6 protocol packets from DHCPv6 clients against link-local address ND snooping entries; checks ND protocol packets against link-local address ND snooping entries, DHCPv6 snooping entries, and static binding entries; and checks the IPv6 data packets from the clients against dynamic binding entries (including link-local address ND snooping entries and DHCPv6 snooping entries) applied on the interfaces connected to the clients and against static binding entries. The items to be examined include MAC address, IPv6 address, VLAN information, and ingress port.

## Configuration procedure

# Enable SAVI.

```
<SwitchB> system-view
[SwitchB] ipv6 savi strict
```

# Enable IPv6.

```
[SwitchB] ipv6
```

# Globally enable DHCPv6 snooping.

```
[SwitchB] ipv6 dhcp snooping enable
```

# Assign interfaces GigabitEthernet 1/0/1, GigabitEthernet 1/0/2, and GigabitEthernet 1/0/3 to VLAN 2.

```
[SwitchB] vlan 2
[SwitchB-vlan2] port gigabitethernet 1/0/1 gigabitethernet 1/0/2 gigabitethernet 1/0/3
```

# Enable DHCPv6 snooping in VLAN 2.

```
[SwitchB-vlan2] ipv6 dhcp snooping vlan enable
[SwitchB] quit
```

# Configure interface GigabitEthernet 1/0/1 as a DHCP snooping trusted port.

```
[SwitchB] interface gigabitethernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] ipv6 dhcp snooping trust
[SwitchB-GigabitEthernet1/0/1] quit
```

# Enable link-local address ND snooping and ND detection.

```
[SwitchB] ipv6 nd snooping enable link-local
[SwitchB] vlan 2
[SwitchB-vlan2] ipv6 nd snooping enable
[SwitchB-vlan2] ipv6 nd detection enable
[SwitchB-vlan2] quit
```

# Configure the dynamic IPv6 source guard binding function on downlink ports GigabitEthernet 1/0/2 and GigabitEthernet 1/0/3.

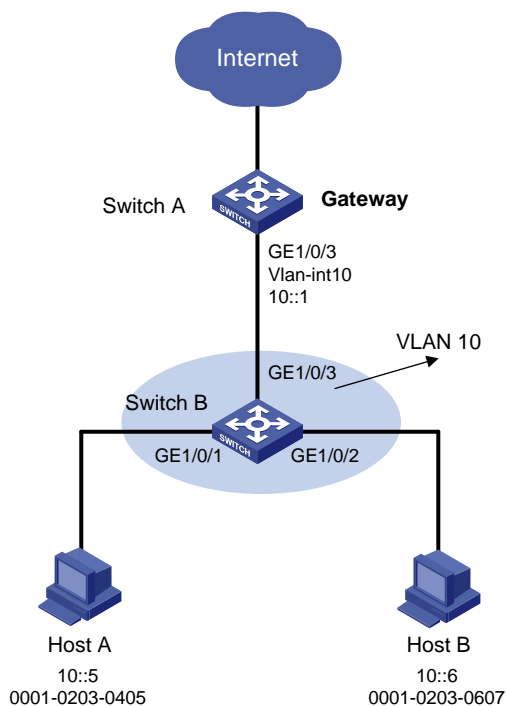
```
[SwitchB] interface gigabitethernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] ipv6 verify source ipv6-address mac-address
[SwitchB-GigabitEthernet1/0/2] quit
[SwitchB] interface gigabitethernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] ipv6 verify source ipv6-address mac-address
[SwitchB-GigabitEthernet1/0/3] quit
```



# SAVI configuration in SLAAC-only address assignment scenario

## Network requirements

Figure 141 Network diagram



As shown in [Figure 141](#), Switch A serves as the gateway. Switch B connects Host A and Host B. The hosts can obtain IPv6 addresses only through SLAAC. Configure SAVI on Switch B to bind the addresses assigned through SLAAC and permit only packets from the bound addresses.

## Configuration considerations

Configure Switch B as follows:

- Enable SAVI.
- Enable global unicast address ND snooping and link-local address ND snooping. For more information about ND snooping, see *Layer 3—IP Services Configuration Guide*.
- Enable ND detection in VLAN 10 to check the ND packets arrived on the ports. For more information about ND detection, see "[Configuring ND attack defense](#)."
- Configure a static IPv6 source guard binding entry on each interface connected to a host. This step is optional. If this step is not performed, SAVI does not check packets against static binding entries. For more information about static IPv6 source guard binding entries, see "[Configuring IP source guard](#)."
- Configure dynamic IPv6 source guard binding on the interfaces connected to the hosts. For more information about dynamic IPv6 source guard binding, see "[Configuring IP source guard](#)."
- Enable DHCPv6 snooping and leave the interface connected to the gateway as its default status (non-trusted port) so that the hosts cannot obtain IP addresses through DHCPv6. For more information about DHCPv6 snooping, see *Layer 3—IP Services Configuration Guide*.

## Packet check principles

Switch B checks ND protocol packets against ND snooping entries and static binding entries; and checks the IPv6 data packets from the hosts against dynamic binding entries (including ND snooping entries) applied on the interfaces connected to the hosts and against static binding entries. The items to be examined include MAC address, IPv6 address, VLAN information, and ingress port.

## Configuration procedure

# Enable SAVI.

```
<SwitchB> system-view
[SwitchB] ipv6 savi strict
```

# Enable IPv6.

```
[SwitchB] ipv6
```

# Assign GigabitEthernet 1/0/1, GigabitEthernet 1/0/2, and GigabitEthernet 1/0/3 to VLAN 10.

```
[SwitchB] vlan 10
[SwitchB-vlan10] port gigabitethernet 1/0/1 gigabitethernet 1/0/2 gigabitethernet 1/0/3
[SwitchB-vlan10] quit
```

# Enable global unicast address ND snooping and link-local address ND snooping.

```
[SwitchB] ipv6 nd snooping enable link-local
[SwitchB] ipv6 nd snooping enable global
[SwitchB] vlan 10
[SwitchB-vlan10] ipv6 nd snooping enable
```

# Enable ND detection.

```
[SwitchB-vlan10] ipv6 nd detection enable
[SwitchB-vlan10] quit
```

# Enable DHCPv6 snooping.

```
[SwitchB] ipv6 dhcp snooping enable
```

# Configure uplink port GigabitEthernet 1/0/3 as an ND trusted port.

```
[SwitchB] interface gigabitethernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] ipv6 nd detection trust
[SwitchB-GigabitEthernet1/0/3] quit
```

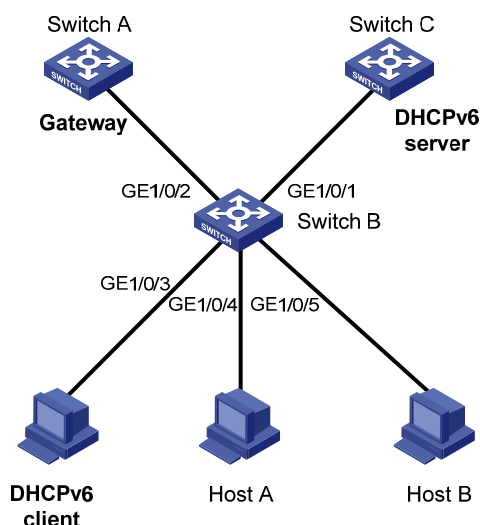
# Configure the dynamic IPv6 source guard binding function on downlink ports GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2.

```
[SwitchB] interface gigabitethernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] ipv6 verify source ipv6-address mac-address
[SwitchB-GigabitEthernet1/0/1] quit
[SwitchB] interface gigabitethernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] ipv6 verify source ipv6-address mac-address
[SwitchB-GigabitEthernet1/0/2] quit
```

# SAVI configuration in DHCPv6+SLAAC address assignment scenario

## Network requirements

Figure 142 Network diagram



As shown in [Figure 142](#), Switch B connects to the DHCPv6 server through interface GigabitEthernet 1/0/1 and connects to the DHCPv6 client through interface GigabitEthernet 1/0/3. Host A and Host B access Gateway (Switch A) through Switch B. Interfaces GigabitEthernet 1/0/1 through GigabitEthernet 1/0/5 on Switch B belong to VLAN 2. The hosts can obtain IP addresses through DHCPv6 or SLAAC. Configure SAVI on Switch B to permit only packets from addresses assigned through DHCPv6 and the bound addresses assigned through SLAAC.

## Configuration considerations

Configure Switch B as follows:

- Enable SAVI.
- Enable DHCPv6 snooping. For more information about DHCPv6 snooping, see *Layer 3—IP Services Configuration Guide*.
- Enable global unicast address ND snooping and link-local address ND snooping. For more information about ND snooping, see *Layer 3—IP Services Configuration Guide*.
- Enable ND detection in VLAN 2 to check the ND packets arrived on the ports. For more information about ND detection, see "[Configuring ND attack defense](#)."
- Configure a static IPv6 source guard binding entry on each interface connected to a host. This step is optional. If this step is not performed, SAVI does not check packets against static binding entries. For more information about static IPv6 source guard binding entries, see "[Configuring IP source guard](#)."
- Configure dynamic IPv6 source guard binding on the interfaces connected to the hosts. For more information about dynamic IPv6 source guard binding, see "[Configuring IP source guard](#)."

## Packet check principles

Switch B checks DHCPv6 protocol packets from DHCPv6 clients against link-local address ND snooping entries; checks ND protocol packets against ND snooping entries, DHCPv6 snooping entries, and static

binding entries; and checks the IPv6 data packets from the hosts against dynamic binding entries (including ND snooping entries and DHCPv6 snooping entries) applied on the interfaces connected to the hosts and against static binding entries. The items to be examined include MAC address, IPv6 address, VLAN information, and ingress port.

## Configuration procedure

# Enable SAVI.

```
<SwitchB> system-view
[SwitchB] ipv6 savi strict
```

# Enable IPv6.

```
[SwitchB] ipv6
```

# Enable DHCPv6 snooping.

```
[SwitchB] ipv6 dhcp snooping enable
```

# Assign interfaces GigabitEthernet 1/0/1 through GigabitEthernet 1/0/5 to VLAN 2.

```
[SwitchB] vlan 2
[SwitchB-vlan2] port gigabitethernet 1/0/1 gigabitethernet 1/0/2 gigabitethernet 1/0/3
gigabitethernet 1/0/4 gigabitethernet 1/0/5
```

# Enable DHCPv6 snooping in VLAN 2.

```
[SwitchB-vlan2] ipv6 dhcp snooping vlan enable
[SwitchB] quit
```

# Configure interface GigabitEthernet 1/0/1 as a DHCPv6 snooping trusted port.

```
[SwitchB] interface gigabitethernet 1/0/1
[SwitchB-GigabitEthernet1/0/1] ipv6 dhcp snooping trust
[SwitchB-GigabitEthernet1/0/1] quit
```

# Enable ND snooping and ND detection.

```
[SwitchB] ipv6 nd snooping enable link-local
[SwitchB] ipv6 nd snooping enable global
[SwitchB] vlan 2
[SwitchB-vlan2] ipv6 nd snooping enable
[SwitchB-vlan2] ipv6 nd detection enable
[SwitchB-vlan2] quit
```

# Configure interface GigabitEthernet 1/0/2 as an ND detection trusted port.

```
[SwitchB] interface gigabitethernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] ipv6 nd detection trust
[SwitchB-GigabitEthernet1/0/2] quit
```

# Configure the dynamic IPv6 source guard binding function on downlink ports GigabitEthernet 1/0/3 through GigabitEthernet 1/0/5.

```
[SwitchB] interface gigabitethernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] ipv6 verify source ipv6-address mac-address
[SwitchB-GigabitEthernet1/0/3] quit
[SwitchB] interface gigabitethernet 1/0/4
[SwitchB-GigabitEthernet1/0/4] ipv6 verify source ipv6-address mac-address
[SwitchB-GigabitEthernet1/0/4] quit
[SwitchB] interface gigabitethernet 1/0/5
[SwitchB-GigabitEthernet1/0/5] ipv6 verify source ipv6-address mac-address
```

# Configuring blacklist

## Overview

The blacklist feature is an attack prevention mechanism that filters packets based on the source IP address. Compared with ACL-based packet filtering, the blacklist feature is easier to configure and fast in filtering packets sourced from particular IP addresses.

The device can dynamically add and remove blacklist entries by cooperating with the login user authentication feature. When the device detects that a user tried to use FTP, Telnet, SSH, SSL, or web to log in to the device for a specific number of times but failed to log in, it considers the user an invalid user and automatically blacklists the user's IP address to filter subsequent packets sourced from that IP address. This function can effectively prevent users from cracking passwords by repeatedly trying to log in.

The device always uses the login failure threshold of 6 and sets the aging time of a dynamic blacklist entry to 10 minutes. These two settings are not configurable. User login failure reasons include wrong username, wrong password, and wrong verification code (for web users).

The device also supports adding and removing blacklist entries manually. Manually configured blacklist entries fall into two categories: permanent and non-permanent. A permanent blacklist entry is always present unless being removed manually, whereas a non-permanent blacklist entry has a limited lifetime depending on your configuration. When the lifetime of a non-permanent entry expires, the device removes the entry from the blacklist, allowing the packets of the IP address defined by the entry to pass through.

## Configuring the blacklist feature

| Step                             | Command                                                                           | Remarks                                                                                         |
|----------------------------------|-----------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|
| 1. Enter system view.            | <b>system-view</b>                                                                | N/A                                                                                             |
| 2. Enable the blacklist feature. | <b>blacklist enable</b>                                                           | Disabled by default.                                                                            |
| 3. Add a blacklist entry.        | <b>blacklist ip</b> <i>source-ip-address</i><br>[ <b>timeout</b> <i>minutes</i> ] | Optional.<br>To add a permanent entry, do not specify the <b>timeout</b> <i>minutes</i> option. |

## Displaying and maintaining the blacklist

| Task                           | Command                                                                                                                                                                                                                             | Remarks               |
|--------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Display blacklist information. | <b>display blacklist</b> { <b>all</b>   <b>ip</b> <i>source-ip-address</i> [ <b>slot</b> <i>slot-number</i> ]   <b>slot</b> <i>slot-number</i> } [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] | Available in any view |

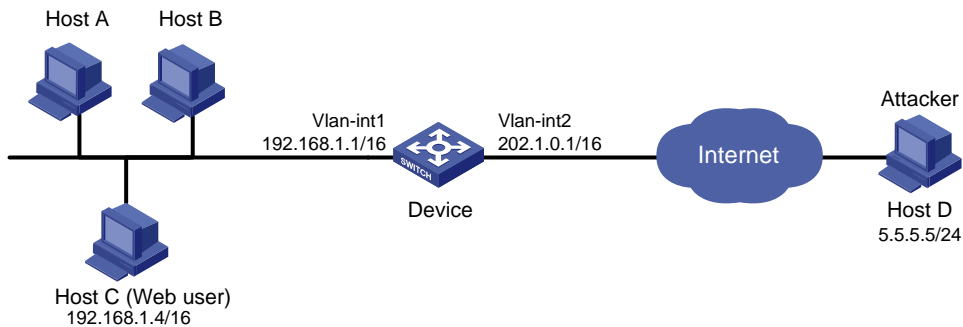
# Blacklist configuration example

## Network requirements

As shown in Figure 143, Host A, Host B, and Host C are internal users, and external user Host D is considered an attacker.

Configure Device to always filter packets from Host D, and to prevent internal users from guessing passwords.

Figure 143 Network diagram



## Configuration procedure

# Assign IP addresses to the interfaces of Device. (Details not shown.)

# Enable the blacklist feature.

```
<Device> system-view
[Device] blacklist enable
```

# Add the IP address of Host D 5.5.5.5 to the blacklist. Do not specify any aging time to make the entry never age out.

```
[Device] blacklist ip 5.5.5.5
```

## Verifying the configuration

If Host C tries to log in to Device through web for six times but fails to log in, the device blacklists Host C. Use the **display blacklist all** command to view all added blacklist entries.

```
[Device] display blacklist all
```

```
Blacklist information

Blacklist : enabled
Blacklist items : 2

IP Type Aging started Aging finished Dropped packets
 YYYY/MM/DD hh:mm:ss YYYY/MM/DD hh:mm:ss
5.5.5.5 manual 2011/04/09 16:02:20 Never 0
192.168.1.4 manual 2011/04/09 16:02:26 2011/04/09 16:12:26 0
```

Host D and Host C are on the blacklist. Host C will stay on the list for 10 minutes, and will then be able to try to log in again. The entry for Host D will never age out. When you do not consider Host D an attacker anymore, you can use the **undo blacklist ip 5.5.5.5** command to remove the entry.

# Index

## [A](#) [B](#) [C](#) [D](#) [E](#) [H](#) [I](#) [L](#) [M](#) [N](#) [O](#) [P](#) [R](#) [S](#) [T](#) [U](#)

### A

- AAA configuration considerations and task list, [15](#)
- AAA configuration examples, [50](#)
- AAA overview, [1](#)
- Applying a QoS policy, [228](#)
- ARP attack protection configuration task list, [351](#)

### B

- Basic configuration for MAC authentication, [118](#)
- Blacklist configuration example, [389](#)

### C

- Configuration prerequisites, [111](#)
- Configuration prerequisites, [92](#)
- Configuration prerequisites, [141](#)
- Configuration task list, [335](#)
- Configuration task list, [118](#)
- Configuration task list, [326](#)
- Configuration task list, [208](#)
- Configuration task list, [245](#)
- Configuring a free IP, [111](#)
- Configuring a MAC authentication critical VLAN, [121](#)
- Configuring a MAC authentication guest VLAN, [120](#)
- Configuring a NAS ID-VLAN binding, [47](#)
- Configuring a PKI domain, [258](#)
- Configuring a switch as a RADIUS server, [48](#)
- Configuring AAA methods for ISP domains, [40](#)
- Configuring AAA schemes, [16](#)
- Configuring an 802.1X critical VLAN, [102](#)
- Configuring an 802.1X guest VLAN, [100](#)
- Configuring an access control policy, [264](#)
- Configuring an Auth-Fail VLAN, [101](#)
- Configuring an entity DN, [257](#)
- Configuring an SSL client policy, [329](#)
- Configuring an SSL server policy, [326](#)
- Configuring ARP active acknowledgement, [358](#)
- Configuring ARP automatic scanning and fixed ARP, [366](#)

- Configuring ARP defense against IP packet attacks, [352](#)
- Configuring ARP detection, [359](#)
- Configuring ARP filtering, [368](#)
- Configuring ARP gateway protection, [367](#)
- Configuring ARP packet rate limit, [355](#)
- Configuring ARP packet source MAC address consistency check, [358](#)
- Configuring global SAVI, [381](#)
- Configuring HABP, [241](#)
- Configuring IPsec for IPv6 routing protocols, [278](#)
- Configuring password control, [233](#)
- Configuring PKI certificate verification, [262](#)
- Configuring port security features, [211](#)
- Configuring portal detection functions, [157](#)
- Configuring portal stateful failover (available only on the HP 5500 EI series), [155](#)
- Configuring RADIUS related attributes, [153](#)
- Configuring secure MAC addresses, [212](#)
- Configuring source MAC address based ARP attack detection, [356](#)
- Configuring the authentication trigger function, [97](#)
- Configuring the blacklist feature, [388](#)
- Configuring the IPv4 source guard function, [336](#)
- Configuring the IPv6 source guard function, [338](#)
- Configuring the local portal server, [143](#)
- Configuring the ND detection function, [372](#)
- Configuring the online user handshake function, [96](#)
- Configuring the quiet timer, [99](#)
- Configuring the redirect URL, [112](#)
- Configuring the switch as an SCP server, [321](#)
- Configuring the switch as an SFTP client, [311](#)
- Configuring the switch as an SFTP server, [310](#)
- Configuring the switch as an SSH client, [294](#)
- Configuring the switch as an SSH server, [289](#)
- Configuring the switch as the SCP client, [321](#)
- Configuring triple authentication, [196](#)
- Configuring URPF, [379](#)



Controlled/uncontrolled port and port authorization status, [78](#)

Controlling access of portal users, [149](#)

Creating a local asymmetric key pair, [246](#)

Creating a user profile, [227](#)

## D

Deleting a certificate, [263](#)

Destroying a local asymmetric key pair, [248](#)

Destroying a local RSA key pair, [263](#)

Displaying and maintaining 802.1X, [103](#)

Displaying and maintaining AAA, [50](#)

Displaying and maintaining EAD fast deployment, [112](#)

Displaying and maintaining HABP, [242](#)

Displaying and maintaining IP source guard, [341](#)

Displaying and maintaining IPsec, [281](#)

Displaying and maintaining MAC authentication, [122](#)

Displaying and maintaining password control, [236](#)

Displaying and maintaining PKI, [264](#)

Displaying and maintaining port security, [214](#)

Displaying and maintaining portal, [160](#)

Displaying and maintaining public keys, [249](#)

Displaying and maintaining SSH, [296](#)

Displaying and maintaining SSL, [330](#)

Displaying and maintaining TCP attack protection, [332](#)

Displaying and maintaining the blacklist, [388](#)

Displaying and maintaining user profiles, [229](#)

Displaying or exporting the local host public key, [246](#)

## E

EAD fast deployment configuration example, [113](#)

Enabling 802.1X, [93](#)

Enabling a user profile, [228](#)

Enabling EAP relay or EAP termination, [93](#)

Enabling port security, [209](#)

Enabling portal authentication, [147](#)

Enabling source MAC consistency check for ND packets, [372](#)

Enabling the periodic online user re-authentication function, [99](#)

Enabling the SYN Cookie feature, [332](#)

## H

HP implementation of 802.1X, [87](#)

HABP configuration example, [242](#)

HABP overview, [240](#)

## I

Ignoring authorization information from the server, [214](#)

Initiating 802.1X authentication, [81](#)

IP source guard configuration examples, [341](#)

IPsec for RIPng configuration example, [281](#)

## L

Logging off portal users, [160](#)

## M

MAC authentication configuration examples, [122](#)

MAC authentication overview, [116](#)

## N

ND detection configuration example, [374](#)

## O

Overview, [245](#)

Overview, [388](#)

Overview, [321](#)

Overview, [129](#)

Overview, [254](#)

Overview, [332](#)

Overview, [195](#)

Overview, [371](#)

Overview, [351](#)

Overview, [310](#)

Overview, [205](#)

Overview, [325](#)

Overview, [286](#)

Overview, [334](#)

Overview, [275](#)

Overview, [111](#)

## P

Password control configuration example, [237](#)

Password control configuration task list, [232](#)

Password control overview, [230](#)

PKI configuration examples, [265](#)

PKI configuration task list, [256](#)

Port security configuration examples, [215](#)

Portal configuration examples, [161](#)

Portal configuration task list, [140](#)

Public key configuration examples, [249](#)

## R

Retrieving a certificate manually, [261](#)

## S

- SAVI configuration in DHCPv6+SLAAC address assignment scenario, [386](#)
- SAVI configuration in DHCPv6-only address assignment scenario, [382](#)
- SAVI configuration in SLAAC-only address assignment scenario, [384](#)
- SAVI overview, [381](#)
- Setting port security's limit on the number of MAC addresses on a port, [209](#)
- Setting the 802.1X authentication timeout timers, [96](#)
- Setting the EAD rule timer, [112](#)
- Setting the maximum number of authentication request attempts, [96](#)
- Setting the maximum number of concurrent 802.1X users on a port, [95](#)
- Setting the port authorization state, [94](#)
- Setting the port security mode, [210](#)
- SFTP client configuration example, [314](#)
- SFTP server configuration example, [318](#)
- Specifying a MAC authentication domain, [120](#)
- Specifying a mandatory authentication domain on a port, [98](#)
- Specifying a source IP address for outgoing portal packets, [154](#)
- Specifying an access control method, [95](#)
- Specifying an Auth-Fail VLAN for portal authentication, [152](#)

- Specifying an auto redirection URL for authenticated portal users, [156](#)
- Specifying supported domain name delimiters, [103](#)
- Specifying the device ID used in stateful failover mode, [48](#)
- Specifying the peer public key on the local device, [248](#)
- Specifying the portal server, [142](#)
- SSH client configuration examples, [304](#)
- SSH server configuration examples, [297](#)
- Submitting a PKI certificate request, [259](#)

## T

- Tearing down user connections, [47](#)
- Triple authentication configuration examples, [197](#)
- Troubleshooting AAA, [76](#)
- Troubleshooting EAD fast deployment, [115](#)
- Troubleshooting IP source guard, [350](#)
- Troubleshooting PKI, [273](#)
- Troubleshooting port security, [224](#)
- Troubleshooting portal, [193](#)
- Troubleshooting SSL, [330](#)

## U

- URPF configuration example, [379](#)
- URPF overview, [376](#)
- User profile configuration task list, [227](#)
- User profile overview, [227](#)
- Using MAC authentication with other features, [117](#)

---

# Contents

|                                                                       |    |
|-----------------------------------------------------------------------|----|
| High availability overview                                            | 1  |
| Availability requirements                                             | 1  |
| Availability evaluation                                               | 1  |
| High availability technologies                                        | 2  |
| Fault detection technologies                                          | 2  |
| Protection switchover technologies                                    | 3  |
| Configuring Ethernet OAM                                              | 5  |
| Ethernet OAM overview                                                 | 5  |
| Major functions of Ethernet OAM                                       | 5  |
| Ethernet OAMPDUs                                                      | 5  |
| How Ethernet OAM works                                                | 6  |
| Standards and protocols                                               | 8  |
| Ethernet OAM configuration task list                                  | 8  |
| Configuring basic Ethernet OAM functions                              | 9  |
| Configuring the Ethernet OAM connection detection timers              | 9  |
| Configuring link monitoring                                           | 10 |
| Configuring errored symbol event detection                            | 10 |
| Configuring errored frame event detection                             | 10 |
| Configuring errored frame period event detection                      | 10 |
| Configuring errored frame seconds event detection                     | 11 |
| Configuring Ethernet OAM remote loopback                              | 11 |
| Enabling Ethernet OAM remote loopback                                 | 11 |
| Rejecting the Ethernet OAM remote loopback request from a remote port | 12 |
| Displaying and maintaining Ethernet OAM configuration                 | 13 |
| Ethernet OAM configuration example                                    | 13 |
| Configuring CFD                                                       | 16 |
| CFD overview                                                          | 16 |
| Basic concepts in CFD                                                 | 16 |
| CFD functions                                                         | 18 |
| Protocols and standards                                               | 20 |
| CFD configuration task list                                           | 20 |
| Configuring basic CFD settings                                        | 21 |
| Enabling CFD                                                          | 21 |
| Configuring the CFD protocol version                                  | 21 |
| Configuring service instances                                         | 22 |
| Configuring MEPs                                                      | 22 |
| Configuring MIP generation rules                                      | 23 |
| Configuring CFD functions                                             | 24 |
| Configuration prerequisites                                           | 24 |
| Configuring CC on MEPs                                                | 24 |
| Configuring LB on MEPs                                                | 25 |
| Configuring LT on MEPs                                                | 25 |
| Configuring AIS                                                       | 26 |
| Configuring LM                                                        | 26 |
| Configuring one-way DM                                                | 26 |
| Configuring two-way DM                                                | 27 |
| Configuring TST                                                       | 27 |
| Displaying and maintaining CFD                                        | 28 |

|                                                                      |           |
|----------------------------------------------------------------------|-----------|
| CFD configuration example .....                                      | 29        |
| <b>Configuring DLDP .....</b>                                        | <b>35</b> |
| DLDP overview .....                                                  | 35        |
| Background .....                                                     | 35        |
| How DLDP works .....                                                 | 36        |
| DLDP configuration task list .....                                   | 42        |
| Configuring the duplex mode and speed of an Ethernet interface ..... | 42        |
| Enabling DLDP .....                                                  | 43        |
| Setting DLDP mode .....                                              | 43        |
| Setting the interval to send advertisement packets .....             | 43        |
| Setting the delaydown timer .....                                    | 44        |
| Setting the port shutdown mode .....                                 | 44        |
| Configuring DLDP authentication .....                                | 45        |
| Resetting DLDP state .....                                           | 45        |
| Displaying and maintaining DLDP .....                                | 46        |
| DLDP configuration examples .....                                    | 46        |
| Automatically shutting down unidirectional links .....               | 46        |
| Manually shutting down unidirectional links .....                    | 50        |
| Troubleshooting DLDP .....                                           | 53        |
| <b>Configuring RRPP .....</b>                                        | <b>54</b> |
| RRPP overview .....                                                  | 54        |
| Background .....                                                     | 54        |
| Basic concepts in RRPP .....                                         | 54        |
| RRPPDUS .....                                                        | 56        |
| RRPP timers .....                                                    | 57        |
| How RRPP works .....                                                 | 57        |
| Typical RRPP networking .....                                        | 59        |
| Protocols and standards .....                                        | 62        |
| RRPP configuration task list .....                                   | 62        |
| Creating an RRPP domain .....                                        | 63        |
| Configuring control VLANs .....                                      | 63        |
| Configuration guidelines .....                                       | 63        |
| Configuration procedure .....                                        | 63        |
| Configuring protected VLANs .....                                    | 64        |
| Configuring RRPP rings .....                                         | 65        |
| Configuring RRPP ports .....                                         | 65        |
| Configuring RRPP nodes .....                                         | 66        |
| Activating an RRPP domain .....                                      | 67        |
| Configuring RRPP timers .....                                        | 68        |
| Configuring an RRPP ring group .....                                 | 68        |
| Configuration restrictions and guidelines .....                      | 68        |
| Configuration procedure .....                                        | 69        |
| Displaying and maintaining RRPP .....                                | 69        |
| RRPP configuration examples .....                                    | 69        |
| Single ring configuration example .....                              | 69        |
| Intersecting ring configuration example .....                        | 72        |
| Dual homed rings configuration example .....                         | 77        |
| Intersecting-ring load balancing configuration example .....         | 87        |
| Troubleshooting .....                                                | 96        |
| <b>Configuring Smart Link .....</b>                                  | <b>97</b> |
| Smart Link overview .....                                            | 97        |
| Background .....                                                     | 97        |
| Terminology .....                                                    | 98        |

|                                                                           |            |
|---------------------------------------------------------------------------|------------|
| How Smart Link works .....                                                | 99         |
| Smart Link collaboration mechanisms .....                                 | 99         |
| Smart Link configuration task list .....                                  | 100        |
| Configuring a Smart Link device .....                                     | 100        |
| Configuration prerequisites .....                                         | 100        |
| Configuring protected VLANs for a smart link group .....                  | 101        |
| Configuring member ports for a smart link group .....                     | 102        |
| Configuring role preemption for a smart link group .....                  | 102        |
| Enabling the sending of flush messages .....                              | 103        |
| Configuring the collaboration between Smart Link and CC of CFD .....      | 103        |
| Configuring an associated device .....                                    | 104        |
| Configuration prerequisites .....                                         | 104        |
| Enabling the receiving of flush messages .....                            | 104        |
| Displaying and maintaining Smart Link .....                               | 104        |
| Smart Link configuration examples .....                                   | 105        |
| Single smart link group configuration example .....                       | 105        |
| Multiple smart link groups load sharing configuration example .....       | 109        |
| Smart Link and CFD collaboration configuration example .....              | 113        |
| <b>Configuring Monitor Link .....</b>                                     | <b>119</b> |
| Monitor Link overview .....                                               | 119        |
| Terminology .....                                                         | 119        |
| How Monitor Link works .....                                              | 120        |
| Configuring Monitor Link .....                                            | 120        |
| Configuration prerequisites .....                                         | 120        |
| Creating a monitor link group .....                                       | 120        |
| Configuring monitor link group member ports .....                         | 120        |
| Displaying and maintaining Monitor Link .....                             | 121        |
| Monitor Link configuration example .....                                  | 121        |
| <b>Configuring VRRP (available only on the HP 5500 EI) .....</b>          | <b>125</b> |
| VRRP overview .....                                                       | 125        |
| VRRP standard protocol mode .....                                         | 126        |
| Introduction to VRRP group .....                                          | 126        |
| VRRP timers .....                                                         | 127        |
| Packet format .....                                                       | 128        |
| Principles of VRRP .....                                                  | 129        |
| VRRP tracking .....                                                       | 129        |
| VRRP application (taking IPv4-based VRRP for example) .....               | 130        |
| VRRP load balancing mode .....                                            | 131        |
| Overview .....                                                            | 131        |
| Assigning virtual MAC addresses .....                                     | 132        |
| Virtual forwarder .....                                                   | 134        |
| Packet types .....                                                        | 136        |
| Configuring VRRP for IPv4 .....                                           | 136        |
| VRRP for IPv4 configuration task list .....                               | 136        |
| Configuring a VRRP operation mode .....                                   | 137        |
| Specifying the type of MAC addresses mapped to virtual IP addresses ..... | 137        |
| Creating a VRRP group and configuring virtual IP address .....            | 138        |
| Configuring router priority, preemptive mode and tracking function .....  | 140        |
| Configuring VF tracking .....                                             | 141        |
| Configuring VRRP packet attributes .....                                  | 142        |
| Enabling the trap function for VRRP .....                                 | 142        |
| Displaying and maintaining VRRP for IPv4 .....                            | 143        |
| Configuring VRRP for IPv6 .....                                           | 143        |

|                                                                         |            |
|-------------------------------------------------------------------------|------------|
| VRRP for IPv6 configuration task list                                   | 143        |
| Specifying the type of MAC addresses mapped to virtual IPv6 addresses   | 144        |
| Creating a VRRP group and configuring a virtual IPv6 address            | 144        |
| Configuring router priority, preemptive mode and tracking function      | 145        |
| Configuring VF tracking                                                 | 146        |
| Configuring VRRP packet attributes                                      | 147        |
| Displaying and maintaining VRRP for IPv6                                | 148        |
| IPv4-based VRRP configuration examples                                  | 148        |
| Single VRRP group configuration example                                 | 148        |
| VRRP interface tracking configuration example                           | 151        |
| VRRP with multiple VLANs configuration example                          | 155        |
| VRRP load balancing mode configuration example                          | 157        |
| IPv6-based VRRP configuration examples                                  | 167        |
| Single VRRP group configuration example                                 | 167        |
| VRRP interface tracking configuration example                           | 170        |
| VRRP with multiple VLANs configuration example                          | 174        |
| VRRP load balancing mode configuration example                          | 178        |
| Troubleshooting VRRP                                                    | 187        |
| The screen frequently displays error prompts.                           | 187        |
| Multiple masters are present in the same VRRP group.                    | 187        |
| Frequent VRRP state transition.                                         | 187        |
| <b>Configuring stateful failover (available only on the HP 5500 EI)</b> | <b>188</b> |
| Stateful failover overview                                              | 188        |
| Operating procedure                                                     | 188        |
| Stateful failover states                                                | 189        |
| Introduction to stateful failover configuration                         | 189        |
| Enabling stateful failover                                              | 190        |
| Configuring the backup VLAN                                             | 190        |
| Displaying and maintaining stateful failover                            | 191        |
| Stateful failover configuration example                                 | 191        |
| Configuration guidelines                                                | 192        |
| <b>Configuring BFD (available only on the HP 5500 EI)</b>               | <b>193</b> |
| BFD overview                                                            | 193        |
| How BFD works                                                           | 193        |
| BFD packet format                                                       | 195        |
| Supported features                                                      | 197        |
| Protocols and standards                                                 | 197        |
| Configuring BFD basic functions                                         | 198        |
| Configuration prerequisites                                             | 198        |
| Configuration procedure                                                 | 198        |
| Displaying and maintaining BFD                                          | 199        |
| <b>Configuring track</b>                                                | <b>200</b> |
| Track overview                                                          | 200        |
| Introduction to collaboration                                           | 200        |
| Collaboration fundamentals                                              | 200        |
| Collaboration application example                                       | 201        |
| Track configuration task list                                           | 201        |
| Associating the track module with a detection module                    | 202        |
| Associating track with NQA                                              | 202        |
| Associating track with BFD (available only on the HP 5500 EI)           | 203        |
| Associating track with interface management                             | 203        |
| Associating the track module with an application module                 | 204        |
| Associating track with VRRP (available only on the HP 5500 EI)          | 204        |

|                                                                                                                                                         |     |
|---------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Associating track with static routing .....                                                                                                             | 206 |
| Associating track with PBR (available only on the HP 5500 EI) .....                                                                                     | 207 |
| Displaying and maintaining track entries .....                                                                                                          | 208 |
| Track configuration examples .....                                                                                                                      | 208 |
| VRRP-track-NQA collaboration configuration example (the master monitors the uplink) (available only on the HP 5500 EI) .....                            | 208 |
| Configuring BFD for a VRRP backup to monitor the master (available only on the HP 5500 EI) .....                                                        | 212 |
| Configuring BFD for the VRRP master to monitor the uplinks (available only on the HP 5500 EI) .....                                                     | 215 |
| Static routing-track-NQA collaboration configuration example .....                                                                                      | 218 |
| Static routing-Track-BFD collaboration configuration example (available only on the HP 5500 EI) .....                                                   | 223 |
| VRRP-track-interface management collaboration configuration example (the master monitors the uplink interface) (available only on the HP 5500 EI) ..... | 226 |

|             |     |
|-------------|-----|
| Index ..... | 230 |
|-------------|-----|

# High availability overview

Communication interruptions can seriously affect widely-deployed value-added services such as IPTV and video conference. Therefore, the basic network infrastructures must be able to provide high availability.

The following are the effective ways to improve availability:

- Increasing fault tolerance
- Speeding up fault recovery
- Reducing impact of faults on services

## Availability requirements

Availability requirements fall into three levels based on purpose and implementation.

**Table 1 Availability requirements**

| Level | Requirement                                                  | Solution                                                                                                                                                                                                               |
|-------|--------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1     | Decrease system software and hardware faults                 | <ul style="list-style-type: none"><li>• <b>Hardware</b>—Simplifying circuit design, enhancing production techniques, and performing reliability tests.</li><li>• <b>Software</b>—Reliability design and test</li></ul> |
| 2     | Protect system functions from being affected if faults occur | Device and link redundancy and deployment of switchover strategies                                                                                                                                                     |
| 3     | Enable the system to recover as fast as possible             | Performing fault detection, diagnosis, isolation, and recovery technologies                                                                                                                                            |

The level 1 availability requirement should be considered during the design and production process of network devices. Level 2 should be considered during network design. Level 3 should be considered during network deployment, according to the network infrastructure and service characteristics.

## Availability evaluation

Mean Time Between Failures (MTBF) and Mean Time to Repair (MTTR) are used to evaluate the availability of a network.

### MTBF

MTBF is the predicted elapsed time between inherent failures of a system during operation. It is typically in the unit of hours. A higher MTBF means a high availability.

### MTTR

MTTR is the average time required to repair a failed system. MTTR in a broad sense also involves spare parts management and customer services.

MTTR = fault detection time + hardware replacement time + system initialization time + link recovery time + routing time + forwarding recovery time. A smaller value of each item means a smaller MTTR and a higher availability.



# High availability technologies

Increasing MTBF or decreasing MTTR can enhance the availability of a network. The high availability technologies described in this section meet the level 2 and level 3 high availability requirements by decreasing MTTR.

High availability technologies can be classified as fault detection technologies or protection switchover technologies.

## Fault detection technologies

Fault detection technologies enable detection and diagnosis of network faults. CFD, DLDP, and Ethernet OAM are data link layer fault detection technologies. BFD is a generic fault detection technology that can be used at any layer. NQA is used for diagnosis and evaluation of network quality. Monitor Link and Track work along with other high availability technologies to detect faults through a collaboration mechanism.

**Table 2 Fault detection technologies**

| Technology                             | Introduction                                                                                                                                                                                                                                                                                                     | Reference                                                                         |
|----------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|
| CFD                                    | Connectivity Fault Detection (CFD), which conforms to IEEE 802.1ag Connectivity Fault Management (CFM) and ITU-T Y.1731, is an end-to-end per-VLAN link layer Operations, Administration and Maintenance (OAM) mechanism used for link connectivity detection, fault verification, and fault location.           | "Configuring CFD" in <i>High Availability Configuration Guide</i>                 |
| DLDP                                   | The Device link detection protocol (DLDP) deals with unidirectional links that may occur in a network. Upon detecting a unidirectional link, DLDP, as configured, can shut down the related port automatically or prompt users to take actions to avoid network problems.                                        | "Configuring DLDP" in <i>High Availability Configuration Guide</i>                |
| Ethernet OAM                           | As a tool monitoring Layer 2 link status, Ethernet OAM is mainly used to address common link-related issues on the "last mile". You can monitor the status of the point-to-point link between two directly connected devices by enabling Ethernet OAM on them.                                                   | "Configuring Ethernet OAM" in <i>High Availability Configuration Guide</i>        |
| BFD (available only on the HP 5500 EI) | Bidirectional forwarding detection (BFD) provides a single mechanism to quickly detect and monitor the connectivity of links or IP forwarding in networks. To improve network performance, devices must quickly detect communication failures to restore communication through backup paths as soon as possible. | "Configuring BFD" in <i>High Availability Configuration Guide</i>                 |
| NQA                                    | Network Quality Analyzer (NQA) analyzes network performance, services and service quality through sending test packets, and provides you with network performance and service quality parameters such as jitter, TCP connection delay, FTP connection delay and file transfer rate.                              | "Configuring NQA" in <i>Network Management and Monitoring Configuration Guide</i> |
| Monitor Link                           | Monitor Link works together with Layer 2 topology protocols to adapt the up/down state of a downlink port to the state of an uplink port. This feature enables fast link switchover on a downstream device in response to the uplink state change on its upstream device.                                        | "Configuring Monitor Link" in <i>High Availability Configuration Guide</i>        |

| Technology | Introduction                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | Reference                                                           |
|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------|
| Track      | The track module is used to implement collaboration between different modules. The collaboration here involves three parts: the application modules, the track module, and the detection modules. These modules collaborate with one another through collaboration entries. That is, the detection modules trigger the application modules to perform certain operations through the track module. More specifically, the detection modules probe the link status, network performance and so on, and inform the application modules of the detection result through the track module. Once notified of network status changes, the application modules deal with the changes to avoid communication interruption and network performance degradation. | "Configuring track" in <i>High Availability Configuration Guide</i> |

## Protection switchover technologies

Protection switchover technologies aim at recovering network faults. They back up hardware, link, routing, and service information for switchover in case of network faults, to ensure continuity of network services.

**Table 3 Protection switchover technologies**

| Technology                             | Introduction                                                                                                                                                                                                                                                                                                                                                                     | Reference                                                                                        |
|----------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|
| Ethernet Link Aggregation              | Ethernet link aggregation, most often simply called link aggregation, aggregates multiple physical Ethernet links into one logical link to increase link bandwidth beyond the limits of any one single link. This logical link is an aggregate link. It allows for link redundancy because the member physical links can dynamically back up one another.                        | "Configuring Ethernet link aggregation" in <i>Layer 2—LAN Switching Configuration Guide</i>      |
| Smart Link                             | Smart Link is a feature developed to address the slow convergence issue with STP. It provides link redundancy as well as fast convergence in a dual uplink network, allowing the backup link to take over quickly when the primary link fails.                                                                                                                                   | "Configuring Smart Link" in <i>High Availability Configuration Guide</i>                         |
| MSTP                                   | As a Layer 2 management protocol, the Multiple Spanning Tree Protocol (MSTP) eliminates Layer 2 loops by selectively blocking redundant links in a network, and in the mean time, allows for link redundancy.                                                                                                                                                                    | "Configuring spanning tree" in <i>Layer 2—LAN Switching Configuration Guide</i>                  |
| RRPP                                   | The Rapid Ring Protection Protocol (RRPP) is a link layer protocol designed for Ethernet rings. RRPP can prevent broadcast storms caused by data loops when an Ethernet ring is healthy, and rapidly restore the communication paths between the nodes in the event that a link is disconnected on the ring.                                                                     | "Configuring RRPP" in <i>High Availability Configuration Guide</i>                               |
| FRR (available only on the HP 5500 EI) | Fast Reroute (FRR) provides a quick per-link or per-node protection on an LSP. In this approach, once a link or node fails on a path, FRR comes up to reroute the path to a new link or node to bypass the failed link or node. This can happen as fast as 50 milliseconds minimizing data loss. Protocols such as RIP, OSPF, IS-IS, and static routing support this technology. | <i>Layer 3—IP Routing Configuration Guide/Configuration Guide</i> of the corresponding protocols |

| <b>Technology</b>                                    | <b>Introduction</b>                                                                                                                                                                                                                                                                                                                                                                                                                           | <b>Reference</b>                                                                                 |
|------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|
| GR (available only on the HP 5500 EI)                | Graceful Restart (GR) ensures the continuity of packet forwarding when a protocol, such as BGP, IS-IS, OSPF, IPv6 BGP, IPv6 IS-IS, or OSPFv3, restarts or during an active/standby switchover process. It needs other devices to implement routing information backup and recovery.                                                                                                                                                           | <i>Layer 3—IP Routing Configuration Guide/Configuration Guide of the corresponding protocols</i> |
| NSR (available only on the HP 5500 EI)               | Non-stop Routing (NSR) is a new feature used to ensure non-stop data transmission during an active/standby switchover. Devices that have formed an IRF fabric support this feature. It backs up IP forwarding information from the master to the slave. Upon an active/standby switchover, NSR can complete link state recovery and route re-generation without requiring the cooperation of other devices. Only IS-IS supports this feature. | "Configuring IS-IS" in <i>Layer 3—IP Routing Configuration Guide</i>                             |
| Stateful Failover (available only on the HP 5500 EI) | Two devices back up the services of each other to ensure that the services on them are consistent. If one device fails, the other device can take over the services by using VRRP or dynamic routing protocols. Because the other device has already backed up the services, service traffic can pass through the other device, avoiding service interruption.                                                                                | "Configuring stateful failover" in <i>High Availability Configuration Guide</i>                  |
| VRRP (available only on the HP 5500 EI)              | Virtual Router Redundancy Protocol (VRRP) is an error-tolerant protocol that provides highly reliable default links on multicast and broadcast LANs such as Ethernet, avoiding network interruption due to failure of a single link.                                                                                                                                                                                                          | "Configuring VRRP" in <i>High Availability Configuration Guide</i>                               |

A single availability technology cannot solve all problems. Therefore, a combination of availability technologies, chosen on the basis of detailed analysis of network environments and user requirements, should be used to enhance network availability. For example, access-layer devices should be connected to distribution-layer devices over redundant links, and core-layer devices should be fully meshed. Also, network availability should be considered during planning prior to building a network.

# Configuring Ethernet OAM

## Ethernet OAM overview

Ethernet Operation, Administration and Maintenance (OAM) is a tool that monitors Layer 2 link status and addresses common link-related issues on the "last mile." You can use it to monitor the status of the point-to-point link between two directly connected devices.

## Major functions of Ethernet OAM

Ethernet OAM provides the following functions:

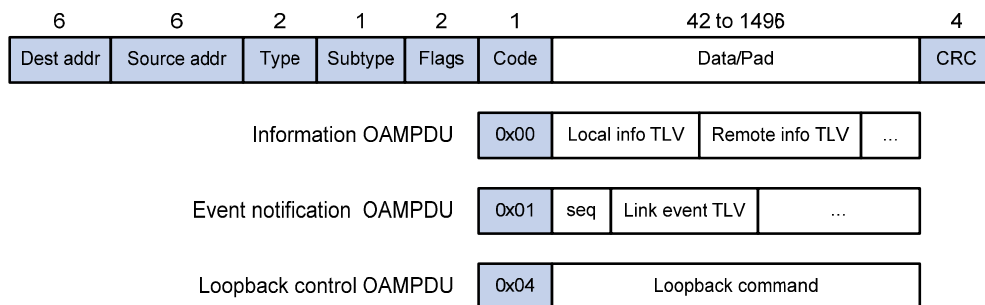
- **Link performance monitoring**—Monitors the performance indices of a link, including packet loss, delay, and jitter, and collects traffic statistics of various types
- **Fault detection and alarm**—Checks the connectivity of a link by sending OAM protocol data units (OAMPDUs) and reports to the network administrators when a link error occurs
- **Remote loopback**—Checks link quality and locates link errors by looping back OAMPDUs

## Ethernet OAMPDUs

Ethernet OAM works on the data link layer. Ethernet OAM reports the link status by periodically exchanging OAMPDUs between devices so that the administrator can effectively manage the network.

Ethernet OAMPDUs fall into the following types: Information, Event Notification, and Loopback Control.

**Figure 1 Formats of different types of Ethernet OAMPDUs**



**Table 4 Fields in an OAMPDU**

| Field       | Description                                                                                                                                                                                     |
|-------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Dest addr   | Destination MAC address of the Ethernet OAMPDU<br>It is a slow protocol multicast address, 0180c2000002. Bridges cannot forward slow protocol packets, so Ethernet OAMPDUs cannot be forwarded. |
| Source addr | Source MAC address of the Ethernet OAMPDU<br>It is the bridge MAC address of the sending side and is a unicast MAC address.                                                                     |
| Type        | Type of the encapsulated protocol in the Ethernet OAMPDU<br>The value is 0x8809.                                                                                                                |

| Field   | Description                                                                           |
|---------|---------------------------------------------------------------------------------------|
| Subtype | The specific protocol being encapsulated in the Ethernet OAMPDU<br>The value is 0x03. |
| Flags   | Status information of an Ethernet OAM entity                                          |
| Code    | Type of the Ethernet OAMPDU                                                           |

**NOTE:**

Throughout this document, a port with Ethernet OAM enabled is an Ethernet OAM entity or an OAM entity.

**Table 5 Functions of different types of OAMPDUs**

| OAMPDU type               | Function                                                                                                                                                                                                                            |
|---------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Information OAMPDU        | Used for transmitting state information of an Ethernet OAM entity—including the information about the local device and remote devices and customized information—to the remote Ethernet OAM entity and maintaining OAM connections. |
| Event Notification OAMPDU | Used by link monitoring to notify the remote OAM entity when it detects problems on the link in between.                                                                                                                            |
| Loopback Control OAMPDU   | Used for remote loopback control. By inserting the information used to enable/disable loopback to a loopback control OAMPDU, you can enable/disable loopback on a remote OAM entity.                                                |

## How Ethernet OAM works

This section describes the working procedures of Ethernet OAM.

### Ethernet OAM connection establishment

Ethernet OAM connection is the basis of all the other Ethernet OAM functions. OAM connection establishment is also known as the "Discovery phase", where an Ethernet OAM entity discovers remote OAM entities and establishes sessions with them.

In this phase, interconnected OAM entities determine whether Ethernet OAM connections can be established, by exchanging Information OAMPDUs to notify the peer of their OAM configuration information and the OAM capabilities of the local nodes. An Ethernet OAM connection can be established between entities that have matching Loopback, link detecting, and link event settings. After an Ethernet OAM connection is established, Ethernet OAM takes effect on both sides.

For Ethernet OAM connection establishment, a device can operate in active Ethernet OAM mode or passive Ethernet OAM mode, but a switch role will be somewhat different depending on the mode.

**Table 6 Active Ethernet OAM mode and passive Ethernet OAM mode**

| Item                             | Active Ethernet OAM mode | Passive Ethernet OAM mode |
|----------------------------------|--------------------------|---------------------------|
| Initiating OAM Discovery         | Available                | Unavailable               |
| Responding to OAM Discovery      | Available                | Available                 |
| Transmitting Information OAMPDUs | Available                | Available                 |

| Item                                            | Active Ethernet OAM mode                           | Passive Ethernet OAM mode |
|-------------------------------------------------|----------------------------------------------------|---------------------------|
| Transmitting Event Notification OAMPDU          | Available                                          | Available                 |
| Transmitting Information OAMPDU without any TLV | Available                                          | Available                 |
| Transmitting Loopback Control OAMPDU            | Available                                          | Unavailable               |
| Responding to Loopback Control OAMPDU           | Available—if both sides operate in active OAM mode | Available                 |

**NOTE:**

- Only OAM entities operating in active OAM mode can initiate OAM connections. OAM entities operating in passive mode wait and respond to the connection requests sent by their peers.
- No OAM connection can be established between OAM entities operating in passive OAM mode.

After an Ethernet OAM connection is established, the Ethernet OAM entities on both sides exchange Information OAMPDU at the handshake packet transmission interval to check whether the Ethernet OAM connection is normal. If an Ethernet OAM entity receives no Information OAMPDU within the Ethernet OAM connection timeout time, the Ethernet OAM connection is considered disconnected.

### Link monitoring

Error detection in an Ethernet is difficult, especially when the physical connection in the network is not disconnected but network performance is degrading gradually. Link monitoring is used to detect and indicate link faults in various environments. Ethernet OAM implements link monitoring through the exchange of Event Notification OAMPDU. When detecting one of the link error events listed in [Table 7](#), the local OAM entity sends an Event Notification OAMPDU to notify the remote OAM entity. With the log information, network administrators can keep track of network status promptly.

**Table 7 Ethernet OAM link error events**

| Ethernet OAM link events    | Description                                                                                                                                                        |
|-----------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Errored symbol event        | An errored symbol event occurs when the number of detected symbol errors during a specified detection interval exceeds the predefined threshold.                   |
| Errored frame event         | An errored frame event occurs when the number of detected error frames during a specified interval exceeds the predefined threshold.                               |
| Errored frame period event  | An errored frame period event occurs if the number of frame errors in a specific number of received frames exceeds the predefined threshold.                       |
| Errored frame seconds event | An errored frame seconds event occurs when the number of error frame seconds detected on a port during a specified detection interval reaches the error threshold. |

The system transforms the period of detecting errored frame period events into the maximum number of 64-byte frames (excluding the interframe spacing and preamble) that a port can send in the specified period. The system takes the maximum number of frames sent as the period. The maximum number of frames sent is calculated using this formula: the maximum number of frames = interface bandwidth (bps) × errored frame period event detection period (in ms)/(64 × 8 × 1000).

A second in which errored frames appear is called an "errored frame second."

## Remote fault detection

Information OAMPDUs are exchanged periodically among Ethernet OAM entities across established OAM connections. In a network where traffic is interrupted due to device failures or unavailability, the flag field defined in information OAMPDUs allows an Ethernet OAM entity to send error information—the critical link event type—to its peer. You can use the log information to track ongoing link status and troubleshoot problems promptly.

**Table 8 Critical link events**

| Type           | Description                                         | OAMPDU transmission frequencies |
|----------------|-----------------------------------------------------|---------------------------------|
| Link Fault     | Peer link signal is lost.                           | Once per second                 |
| Dying Gasp     | A power failure or other unexpected error occurred. | Non-stop                        |
| Critical Event | An undetermined critical event occurred.            | Non-stop                        |

This Switch Series is able to receive information OAMPDUs carrying the critical link events listed in [Table 8](#).

Only the Gigabit fiber ports are able to send information OAMPDUs carrying Link Fault events.

This Switch Series is able to send information OAMPDUs carrying Dying Gasp events when the device is rebooted or relevant ports are manually shut down. Physical IRF ports, however, are unable to send this type of OAMPDU. For more information about physical IRF ports, see *IRF Configuration Guide*.

This Switch Series is unable to send information OAMPDUs carrying Critical Events.

## Remote loopback

Remote loopback is available only after the Ethernet OAM connection is established. With remote loopback enabled, the Ethernet OAM entity operating in active Ethernet OAM mode sends non-OAMPDUs to its peer. After receiving these frames, the peer does not forward them according to their destination addresses. Instead, it returns them to the sender along the original path.

Remote loopback enables you to check the link status and locate link failures. Performing remote loopback periodically helps to detect network faults promptly. Furthermore, performing remote loopback by network segments helps to locate network faults.

## Standards and protocols

Ethernet OAM is defined in IEEE 802.3ah (Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications).

## Ethernet OAM configuration task list

| Task                                                                     | Remarks                                                    |          |
|--------------------------------------------------------------------------|------------------------------------------------------------|----------|
| <a href="#">Configuring basic Ethernet OAM functions</a>                 | Required                                                   |          |
| <a href="#">Configuring the Ethernet OAM connection detection timers</a> | Optional                                                   |          |
| <a href="#">Configuring link monitoring</a>                              | <a href="#">Configuring errored symbol event detection</a> | Optional |
|                                                                          | <a href="#">Configuring errored frame event detection</a>  | Optional |

| Task                                     | Remarks                                                               |
|------------------------------------------|-----------------------------------------------------------------------|
|                                          | Configuring errored frame period event detection                      |
|                                          | Configuring errored frame seconds event detection                     |
| Configuring Ethernet OAM remote loopback | Enabling Ethernet OAM remote loopback                                 |
|                                          | Rejecting the Ethernet OAM remote loopback request from a remote port |

## Configuring basic Ethernet OAM functions

For Ethernet OAM connection establishment, an Ethernet OAM entity operates in active mode or passive mode. Only an Ethernet OAM entity in active mode can initiate connection establishment. After Ethernet OAM is enabled on an Ethernet port, according to its Ethernet OAM mode, the Ethernet port establishes an Ethernet OAM connection with its peer port.

To change the Ethernet OAM mode on an Ethernet OAM-enabled port, you must first disable Ethernet OAM on the port.

To configure basic Ethernet OAM functions:

| Step                                        | Command                                                           | Remarks                                               |
|---------------------------------------------|-------------------------------------------------------------------|-------------------------------------------------------|
| 1. Enter system view.                       | <b>system-view</b>                                                | N/A                                                   |
| 2. Enter Layer 2 Ethernet interface view.   | <b>interface</b> <i>interface-type</i><br><i>interface-number</i> | N/A                                                   |
| 3. Set the Ethernet OAM mode.               | <b>oam mode</b> { <b>active</b>   <b>passive</b> }                | Optional.<br>The default is active Ethernet OAM mode. |
| 4. Enable Ethernet OAM on the current port. | <b>oam enable</b>                                                 | Ethernet OAM is disabled by default.                  |

## Configuring the Ethernet OAM connection detection timers

After an Ethernet OAM connection is established, the Ethernet OAM entities on both sides exchange Information OAMPDUs at the handshake packet transmission interval to check whether the Ethernet OAM connection is normal. If an Ethernet OAM entity receives no Information OAMPDU within the Ethernet OAM connection timeout time, the Ethernet OAM connection is considered disconnected.

By adjusting the handshake packet transmission interval and the connection timeout timer, you can change the detection time resolution for Ethernet OAM connections.

After the timeout timer of an Ethernet OAM connection expires, the local OAM entity ages out its connection with the peer OAM entity, causing the OAM connection to be disconnected. HP recommends that you set the connection timeout timer to at least five times the handshake packet transmission interval, ensuring the stability of Ethernet OAM connections.

To configure the Ethernet OAM connection detection timers:



| Step                                                                  | Command                                    | Remarks                                    |
|-----------------------------------------------------------------------|--------------------------------------------|--------------------------------------------|
| 1. Enter system view.                                                 | <b>system-view</b>                         | N/A                                        |
| 2. Configure the Ethernet OAM handshake packet transmission interval. | <b>oam timer hello</b> <i>interval</i>     | Optional.<br>1000 millisecond by default.  |
| 3. Configure the Ethernet OAM connection timeout timer.               | <b>oam timer keepalive</b> <i>interval</i> | Optional.<br>5000 milliseconds by default. |

## Configuring link monitoring

After Ethernet OAM connections are established, the link monitoring periods and thresholds configured in this section take effect on all Ethernet ports automatically.

## Configuring errored symbol event detection

| Step                                                        | Command                                                    | Remarks                           |
|-------------------------------------------------------------|------------------------------------------------------------|-----------------------------------|
| 1. Enter system view.                                       | <b>system-view</b>                                         | N/A                               |
| 2. Configure the errored symbol event detection interval.   | <b>oam errored-symbol period</b> <i>period-value</i>       | Optional.<br>1 second by default. |
| 3. Configure the errored symbol event triggering threshold. | <b>oam errored-symbol threshold</b> <i>threshold-value</i> | Optional.<br>1 by default.        |

## Configuring errored frame event detection

| Step                                                       | Command                                                   | Remarks                           |
|------------------------------------------------------------|-----------------------------------------------------------|-----------------------------------|
| 1. Enter system view.                                      | <b>system-view</b>                                        | N/A                               |
| 2. Configure the errored frame event detection interval.   | <b>oam errored-frame period</b> <i>period-value</i>       | Optional.<br>1 second by default. |
| 3. Configure the errored frame event triggering threshold. | <b>oam errored-frame threshold</b> <i>threshold-value</i> | Optional.<br>1 by default.        |

## Configuring errored frame period event detection

| Step                                                              | Command                                                          | Remarks                                    |
|-------------------------------------------------------------------|------------------------------------------------------------------|--------------------------------------------|
| 1. Enter system view.                                             | <b>system-view</b>                                               | N/A                                        |
| 2. Configure the errored frame period event detection period.     | <b>oam errored-frame-period period</b> <i>period-value</i>       | Optional.<br>1000 milliseconds by default. |
| 3. Configure the errored frame period event triggering threshold. | <b>oam errored-frame-period threshold</b> <i>threshold-value</i> | Optional.<br>1 by default.                 |

# Configuring errored frame seconds event detection

## IMPORTANT:

Make sure the errored frame seconds triggering threshold is less than the errored frame seconds detection interval. Otherwise, no errored frame seconds event can be generated.

To configure errored frame seconds event detection:

| Step                                                               | Command                                                           | Remarks                            |
|--------------------------------------------------------------------|-------------------------------------------------------------------|------------------------------------|
| 1. Enter system view.                                              | <b>system-view</b>                                                | N/A                                |
| 2. Configure the errored frame seconds event detection interval.   | <b>oam errored-frame-seconds period</b> <i>period-value</i>       | Optional.<br>60 second by default. |
| 3. Configure the errored frame seconds event triggering threshold. | <b>oam errored-frame-seconds threshold</b> <i>threshold-value</i> | Optional.<br>1 by default.         |

# Configuring Ethernet OAM remote loopback

## Enabling Ethernet OAM remote loopback

### ⚠ CAUTION:

Use this function with caution, because enabling Ethernet OAM remote loopback impacts other services.

When you enable Ethernet OAM remote loopback on a port, the port sends Loopback Control OAMPDUs to a remote port, and the remote port enters the loopback state. The port then sends test frames to the remote port. By observing how many of these test frames return, you can calculate the packet loss ratio on the link to evaluate the link performance.

You can enable Ethernet OAM remote loopback on a specific port in user view, system view, or Layer 2 Ethernet interface view. The configuration effects are the same.

### Configuration guidelines

- Ethernet OAM remote loopback is available only after the Ethernet OAM connection is established and can be performed only by Ethernet OAM entities operating in active Ethernet OAM mode.
- Remote loopback is available only on full-duplex links that support remote loopback at both ends.
- Ethernet OAM remote loopback must be supported by both the remote port and the sending port.
- Enabling Ethernet OAM remote loopback interrupts data communications. After Ethernet OAM remote loopback is disabled, all the ports involved will shut down and then come up. Ethernet OAM remote loopback can be disabled by any of the following actions: executing the **undo oam enable** command to disable Ethernet OAM; executing the **undo oam loopback interface** or **undo oam loopback** command to disable Ethernet OAM remote loopback; and Ethernet OAM connection timing out.
- Ethernet OAM remote loopback is only applicable to individual links. It is not applicable to link aggregation member ports or service loopback group member ports. In addition, do not assign ports where Ethernet OAM remote loopback is being performed to link aggregation groups or

service loopback groups. For more information about link aggregation groups and service loopback groups, see *Layer 2—LAN Switching Configuration Guide*.

- Enabling internal loopback test on a port in remote loopback test can terminate the remote loopback test. For more information about loopback test, see *Layer 2—LAN Switching Configuration Guide*.

## Configuration procedure

To enable Ethernet OAM remote loopback in user view:

| Task                                                    | Command                                                                 | Remarks              |
|---------------------------------------------------------|-------------------------------------------------------------------------|----------------------|
| Enable Ethernet OAM remote loopback on a specific port. | <b>oam loopback interface</b><br><i>interface-type interface-number</i> | Disabled by default. |

To enable Ethernet OAM remote loopback in system view:

| Step                                                       | Command                                                                 | Remarks              |
|------------------------------------------------------------|-------------------------------------------------------------------------|----------------------|
| 1. Enter system view.                                      | <b>system-view</b>                                                      | N/A                  |
| 2. Enable Ethernet OAM remote loopback on a specific port. | <b>oam loopback interface</b><br><i>interface-type interface-number</i> | Disabled by default. |

To enable Ethernet OAM remote loopback in Layer 2 Ethernet interface view:

| Step                                                | Command                                                           | Remarks              |
|-----------------------------------------------------|-------------------------------------------------------------------|----------------------|
| 1. Enter system view.                               | <b>system-view</b>                                                | N/A                  |
| 2. Enter Layer 2 Ethernet interface view.           | <b>interface</b> <i>interface-type</i><br><i>interface-number</i> | N/A                  |
| 3. Enable Ethernet OAM remote loopback on the port. | <b>oam loopback</b>                                               | Disabled by default. |

## Rejecting the Ethernet OAM remote loopback request from a remote port

The Ethernet OAM remote loopback function impacts other services. To solve this problem, you can disable a port from being controlled by the Loopback Control OAMPDUs sent by a remote port. The local port then rejects the Ethernet OAM remote loopback request from the remote port.

To reject the Ethernet OAM remote loopback request from a remote port:

| Step                                                                   | Command                                                           | Remarks                                                                                         |
|------------------------------------------------------------------------|-------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|
| 1. Enter system view.                                                  | <b>system-view</b>                                                | N/A                                                                                             |
| 2. Enter Layer 2 Ethernet interface view.                              | <b>interface</b> <i>interface-type</i><br><i>interface-number</i> | N/A                                                                                             |
| 3. Reject the Ethernet OAM remote loopback request from a remote port. | <b>oam loopback reject-request</b>                                | By default, a port does not reject the Ethernet OAM remote loopback request from a remote port. |

# Displaying and maintaining Ethernet OAM configuration

| Task                                                                                                      | Command                                                                                                                                                                                                       | Remarks                |
|-----------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|
| Display global Ethernet OAM configuration.                                                                | <b>display oam configuration</b> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                                                                                           | Available in any view  |
| Display the statistics on critical events after an Ethernet OAM connection is established.                | <b>display oam critical-event</b> [ <b>interface</b> <i>interface-type interface-number</i> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                              | Available in any view  |
| Display the statistics on Ethernet OAM link error events after an Ethernet OAM connection is established. | <b>display oam link-event</b> { <b>local</b>   <b>remote</b> } [ <b>interface</b> <i>interface-type interface-number</i> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] | Available in any view  |
| Display the information about an Ethernet OAM connection.                                                 | <b>display oam</b> { <b>local</b>   <b>remote</b> } [ <b>interface</b> <i>interface-type interface-number</i> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]            | Available in any view  |
| Clear statistics on Ethernet OAM packets and Ethernet OAM link error events.                              | <b>reset oam</b> [ <b>interface</b> <i>interface-type interface-number</i> ]                                                                                                                                  | Available in user view |

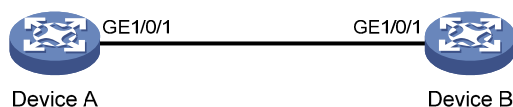
## Ethernet OAM configuration example

### Network requirements

On the network shown in [Figure 2](#), perform the following operations:

- Enable Ethernet OAM on Device A and Device B to auto-detect link errors between the two devices
- Monitor the performance of the link between Device A and Device B by collecting statistics about the error frames received by Device A

**Figure 2 Network diagram**



### Configuration procedure

1. Configure Device A:

```
Configure GigabitEthernet 1/0/1 to operate in passive Ethernet OAM mode and enable Ethernet OAM for it.
```

```
<DeviceA> system-view
```

```
[DeviceA] interface gigabitethernet 1/0/1
```

```
[DeviceA-GigabitEthernet1/0/1] oam mode passive
[DeviceA-GigabitEthernet1/0/1] oam enable
[DeviceA-GigabitEthernet1/0/1] quit
```

# Set the errored frame detection interval to 20 seconds and set the errored frame event triggering threshold to 10.

```
[DeviceA] oam errored-frame period 20
[DeviceA] oam errored-frame threshold 10
```

## 2. Configure Device B:

# Configure GigabitEthernet 1/0/1 to operate in active Ethernet OAM mode (the default) and enable Ethernet OAM for it.

```
<DeviceB> system-view
[DeviceB] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] oam mode active
[DeviceB-GigabitEthernet1/0/1] oam enable
[DeviceB-GigabitEthernet1/0/1] quit
```

## 3. Verify the configuration:

Use the **display oam configuration** command to display the Ethernet OAM configuration. For example:

# Display the Ethernet OAM configuration on Device A.

```
[DeviceA] display oam configuration
Configuration of the link event window/threshold :

Errored-symbol Event period(in seconds) : 1
Errored-symbol Event threshold : 1
Errored-frame Event period(in seconds) : 20
Errored-frame Event threshold : 10
Errored-frame-period Event period(in ms) : 1000
Errored-frame-period Event threshold : 1
Errored-frame-seconds Event period(in seconds) : 60
Errored-frame-seconds Event threshold : 1
```

Configuration of the timer :

```

Hello timer(in ms) : 1000
Keepalive timer(in ms) : 5000
```

The output shows that the detection period of errored frame events is 20 seconds, the detection threshold is 10 seconds, and all the other parameters use the default values.

You can use the **display oam critical-event** command to display the statistics of Ethernet OAM critical link events. For example:

# Display the statistics of Ethernet OAM critical link events on all the ports of Device A.

```
[DeviceA] display oam critical-event
Port : GigabitEthernet1/0/1
Link Status : Up
Event statistic :

Link Fault : 0 Dying Gasp : 0 Critical Event : 0
```

The output shows that no critical link event occurred on the link between Device A and Device B. You can use the **display oam link-event** command to display the statistics of Ethernet OAM link error events. For example:

# Display Ethernet OAM link event statistics of the remote end of Device B.

```
[DeviceB] display oam link-event remote
```

```
Port :GigabitEthernet1/0/1
```

```
Link Status :Up
```

```
OAMRemoteErrFrameEvent : (ms = milliseconds)
```

```

Event Time Stamp : 5789 Errored FrameWindow : 10(100ms)
Errored Frame Threshold : 1 Errored Frame : 3
Error Running Total : 35 Event Running Total : 17
```

The output shows that 35 errors occurred since Ethernet OAM was enabled on Device A, 17 of which are caused by error frames. The link is unstable.

# Configuring CFD

## CFD overview

Connectivity Fault Detection (CFD), which conforms to IEEE 802.1ag Connectivity Fault Management (CFM) and ITU-T Y.1731, is an end-to-end per-VLAN link layer Operations, Administration and Maintenance (OAM) mechanism used for link connectivity detection, fault verification, and fault location.

## Basic concepts in CFD

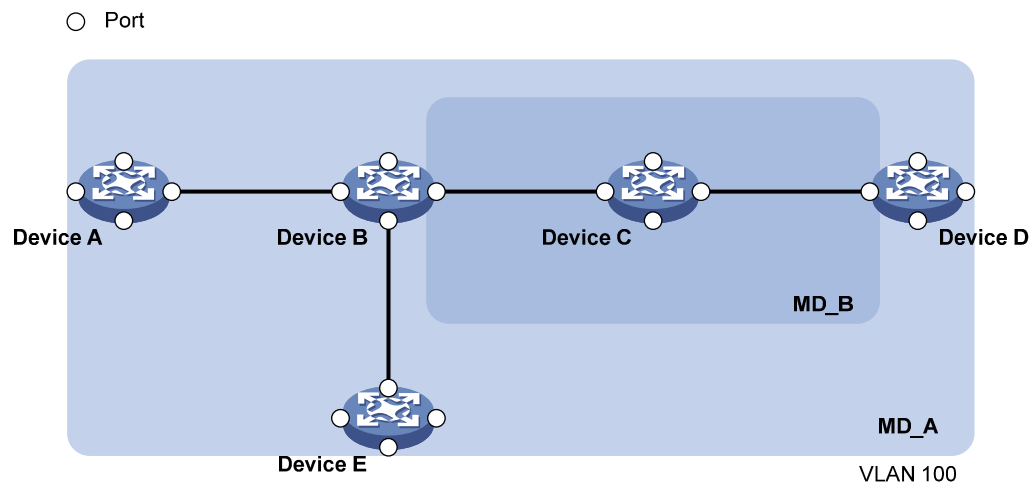
### Maintenance domain

A maintenance domain (MD) defines the network where CFD plays its role. The MD boundary is defined by some maintenance association end points (MEPs) configured on the ports. An MD is identified by its MD name.

To accurately locate faults, CFD introduces eight levels (from 0 to 7) to MDs. The bigger the number, the higher the level and the larger the area covered. Domains can touch or nest (if the outer domain has a higher level than the nested one) but cannot intersect or overlap.

MD levels facilitate fault location and make fault location more accurate. As shown in [Figure 3](#), MD\_A in light blue nests MD\_B in dark blue. If a connectivity fault is detected at the boundary of MD\_A, any of the devices in MD\_A, including Device A through Device E, may fail. If a connectivity fault is also detected at the boundary of MD\_B, the failure points may be any of Device B through Device D. If the devices in MD\_B can operate properly, at least Device C is operational.

**Figure 3 Two nested MDs**



CFD exchanges messages and performs operations on a per-domain basis. By planning MDs properly in a network, you can use CFD to rapidly locate failure points.

### Maintenance association

A maintenance association (MA) is a set of maintenance points (MPs) in an MD. An MA is identified by the "MD name + MA name." You can configure multiple MAs in an MD as needed.

An MA serves a VLAN. Packets sent by the MPs in an MA carry the relevant VLAN tag. An MP can receive packets sent by other MPs in the same MA.

## Maintenance point

An MP is configured on a port and belongs to an MA. MPs fall into two types: maintenance association end points (MEPs) and maintenance association intermediate points (MIPs).

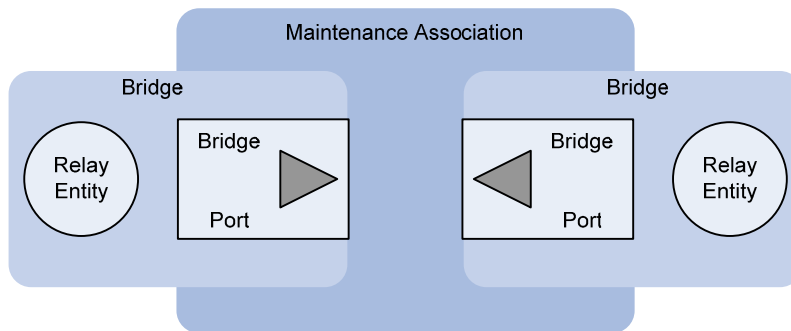
- MEP

Each MEP is identified by an integer called a "MEP ID". The MEPs of an MD define the range and boundary of the MD. The MA and MD that a MEP belongs to define the VLAN attribute and level of the packets sent by the MEP. MEPs fall into inward-facing MEPs and outward-facing MEPs.

The level of a MEP determines the levels of packets that the MEP can process. The packets transmitted from a MEP carry the level of the MEP. A MEP forwards packets at a higher level and processes packet of its level or lower. The processing procedure is specific to packets in the same VLAN. Packets of different VLANs are independent.

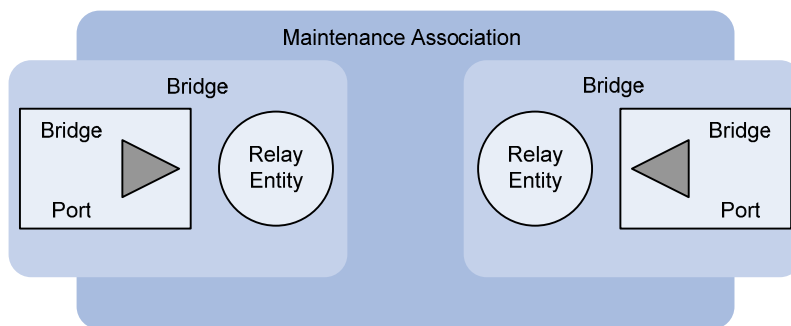
The direction of a MEP (outward-facing or inward-facing) determines the position of the MD relative to the port.

**Figure 4 Outward-facing MEP**



As shown in [Figure 4](#), an outward-facing MEP sends packets to its host port.

**Figure 5 Inward-facing MEP**



As shown in [Figure 5](#), an inward-facing MEP does not send packets to its host port. Rather, it sends packets to other ports on the device.

- MIP

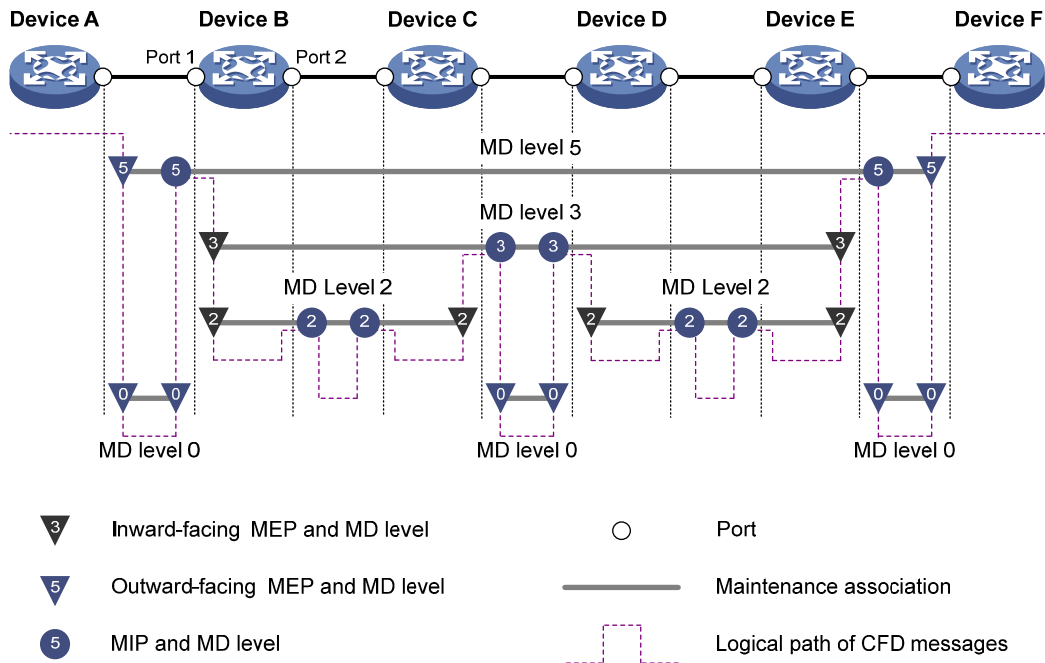
A MIP is internal to an MD. It cannot send CFD packets actively. However, it can handle and respond to CFD packets. The MA and MD to which a MIP belongs define the VLAN attribute and level of the packets received.



By cooperating with MEPs, a MIP can perform a function similar to ping and traceroute. Like a MEP, a MIP forwards packets of a different level without any processing and only processes packets of its level.

Figure 6 demonstrates a grading example of the CFD module. Six devices, labeled A through F respectively, exist. Suppose each device has two ports, and MEPs and MIPs are configured on some of these ports. Four levels of MDs are designed in this example, the bigger the number, the higher the level and the larger the area covered. In this example, Port 1 of device B is configured with the following MPs—a level 5 MIP, a level 3 inward-facing MEP, a level 2 inward-facing MEP, and a level 0 outward-facing MEP.

Figure 6 Levels of MPs



## MEP list

A MEP list is a collection of local MEPs allowed to be configured and the remote MEPs to be monitored in the same MA. It lists all the MEPs configured on different devices in the same MA. The MEPs all have unique MEP IDs. When a MEP receives from a remote device a continuity check message (CCM) that carries a MEP ID not included in the MEP list of the MA, it drops the message.

## CFD functions

CFD works effectively only in properly-configured networks. Its functions, which are implemented through the MPs, include:

- Continuity check (CC)
- Loopback (LB)
- Linktrace (LT)
- Alarm indication signal (AIS)
- Loss measurement (LM)
- Delay measurement (DM)
- Test (TST)

## Continuity check

Connectivity faults are usually caused by device faults or configuration errors. Continuity check checks the connectivity between MEPs. This function is implemented through periodic sending of continuity check messages (CCMs) by the MEPs. As a multicast message, a CCM sent by one MEP is intended to be received by all the other MEPs in the same MA. If a MEP fails to receive the CCMs within 3.5 times the sending interval, the link is considered faulty and a log is generated. When multiple MEPs send CCMs at the same time, the multipoint-to-multipoint link check is achieved. CCM frames are multicast frames.

## Loopback

Similar to ping at the IP layer, loopback verifies the connectivity between a local device and a remote device. To implement this function, the local MEP sends loopback messages (LBMs) to the remote MEP. Depending on whether the local MEP can receive a loopback reply message (LBR) from the remote MEP, the link state between the two can be verified. LBM frames and LBR frames are unicast frames.

## Linktrace

Linktrace identifies the path between the source MEP and the target MEP. This function is implemented in the following way—the source MEP sends the linktrace messages (LTMs) to the target MEP. After receiving the messages, the target MEP and the MIPs that the LTM frames pass send back linktrace reply messages (LTRs) to the source MEP. Based on the reply messages, the source MEP can identify the path to the target MEP. LTM frames are multicast frames and LTRs are unicast frames.

## AIS

The AIS function suppresses the number of error alarms reported by MEPs. If a local MEP receives no CCM frames from its peer MEP within 3.5 times the CCM transmission interval, it immediately starts to send AIS frames periodically in the opposite direction of CCM frames. Upon receiving the AIS frames, the peer MEP suppresses the error alarms locally, and continues to send the AIS frames. If the local MEP receives CCM frames within 3.5 times the CCM transmission interval, it stops sending AIS frames and restores the error alarm function. AIS frames are multicast frames.

## LM

The LM function measures the frame loss in a certain direction between a pair of MEPs. The source MEP sends loss measurement messages (LMMs) to the target MEP, the target MEP responds with loss measurement replies (LMRs), and the source MEP calculates the number of lost frames according to the counter values of the two consecutive LMRs (the current LMR and the previous LMR). LMMs and LMRs are multicast frames.

## DM

The DM function measures frame delays between two MEPs, including one-way and two-way frame delays.

1. One-way frame delay measurement

The source MEP sends a one-way delay measurement (1DM) frame, which carries the transmission time, to the target MEP. Upon receiving the 1DM frame, the target MEP records the reception time, and calculates and records the link transmission delay and jitter (delay variation) according to the transmission time and reception time. 1DM frames are multicast frames.

2. Two-way frame delay measurement

The source MEP sends a delay measurement message (DMM), which carries the transmission time, to the target MEP. Upon receiving the DMM, the target MEP responds with a delay measurement reply (DMR), which carries the reception time and transmission time of the DMM and the transmission time of the DMR. Upon receiving the DMR, the source MEP records the DMR reception

time, and calculates the link transmission delay and jitter according to the DMR reception time and DMM transmission time. DMM frames and DMR frames are multicast frames.

## TST

The TST function tests the bit errors between two MEPs. The source MEP sends a TST frame, which carries the test pattern, such as pseudo random bit sequence (PRBS) or all-zero, to the target MEP. Upon receiving the TST frame, the target MEP determines the bit errors by calculating and comparing the content of the TST frame. TST frames are unicast frames.

## Protocols and standards

- IEEE 802.1ag, *Virtual Bridged Local Area Networks Amendment 5: Connectivity Fault Management*
- ITU-T Y.1731, *OAM functions and mechanisms for Ethernet based networks*

## CFD configuration task list

For CFD to work properly, design the network by performing the following tasks:

- Grade the MDs in the entire network, and define the boundary of each MD
- Assign a name for each MD. Make sure that the same MD has the same name on different devices.
- Define the MA in each MD according to the VLAN you want to monitor
- Assign a name for each MA. Make sure that the same MA in the same MD has the same name on different devices.
- Determine the MEP list of each MA in each MD. Make sure that devices in the same MA maintain the same MEP list.
- At the edges of MD and MA, MEPs should be designed at the device port. MIPs can be designed on devices or ports that are not at the edges.

Complete the following tasks to configure CFD:

| Tasks                          | Remarks                              |                                                                                                                                    |
|--------------------------------|--------------------------------------|------------------------------------------------------------------------------------------------------------------------------------|
|                                | Enabling CFD                         | Required                                                                                                                           |
|                                | Configuring the CFD protocol version | Optional                                                                                                                           |
| Configuring basic CFD settings | Configuring service instances        | Creating a service instance with the MD name<br>Required<br>Creating a service instance without the MD name<br>Perform either task |
|                                | Configuring MEPs                     | Required                                                                                                                           |
|                                | Configuring MIP generation rules     | Required                                                                                                                           |
| Configuring CFD functions      | Configuring CC on MEPs               | Required                                                                                                                           |
|                                | Configuring LB on MEPs               | Optional                                                                                                                           |
|                                | Configuring LT on MEPs               | Optional                                                                                                                           |
|                                | Configuring AIS                      | Optional                                                                                                                           |

| Tasks                                  | Remarks  |
|----------------------------------------|----------|
| <a href="#">Configuring LM</a>         | Optional |
| <a href="#">Configuring one-way DM</a> | Optional |
| <a href="#">Configuring two-way DM</a> | Optional |
| <a href="#">Configuring TST</a>        | Optional |

**NOTE:**

Typically, a port blocked by STP cannot receive or send CFD messages except in the following cases:

- The port is configured as an outward-facing MEP.
- The port is configured as a MIP or inward-facing MEP, which can still receive and send CFD messages except CCM messages.

## Configuring basic CFD settings

### Enabling CFD

Enable CFD on all concerned devices.

To enable CFD on a device:

| Step                  | Command            | Remarks                     |
|-----------------------|--------------------|-----------------------------|
| 1. Enter system view. | <b>system-view</b> | N/A                         |
| 2. Enable CFD.        | <b>cfid enable</b> | CFD is disabled by default. |

### Configuring the CFD protocol version

Three CFD protocol versions are available: IEEE 802.1ag draft5.2 version, IEEE 802.1ag draft5.2 interim version, and IEEE 802.1ag standard version. Devices in a same MD must use the same CFD protocol version. Otherwise, they cannot exchange CFD protocol packets.

To configure the CFD protocol version:

| Step                                   | Command                                                 | Remarks                                                                 |
|----------------------------------------|---------------------------------------------------------|-------------------------------------------------------------------------|
| 1. Enter system view.                  | <b>system-view</b>                                      | N/A                                                                     |
| 2. Configure the CFD protocol version. | <b>cfid version { draft5   draft5-plus   standard }</b> | Optional.<br>By default, CFD uses the standard version of IEEE 802.1ag. |

If an MD is created by using the **cfid md** command or automatically generated by using the **cfid service-instance maid format** command on a device, you cannot switch between the standard and non-standard versions (draft5.2 version and draft5.2 interim version). However, you can switch between the draft5.2 version and draft5.2 interim version. This restriction does not apply to the device without an MD configured.

## Configuring service instances

Before configuring the MEPs and MIPs, you must first configure service instances. A service instance is a set of service access points (SAPs), and it belongs to an MA in an MD.

A service instance is indicated by an integer to represent an MA in an MD. The MD and MA define the level and VLAN attribute of the messages handled by the MPs in a service instance.

Service instances fall into two types:

- Service instance with the MD name, which takes effect in any version of CFD.
- Service instance without the MD name, which takes effect in only CFD IEEE 802.1ag.

You can create either type of service instance as needed.

### Creating a service instance with the MD name

To create a service instance with the MD name, create the MD and MA for the service instance first.

#### CAUTION:

You must create the MD, MA, and service instance by strictly following the order stated in the table.

To configure a service instance with the MD name:

| Step                                           | Command                                                                                           | Remarks                 |
|------------------------------------------------|---------------------------------------------------------------------------------------------------|-------------------------|
| 1. Enter system view.                          | <b>system-view</b>                                                                                | N/A                     |
| 2. Create an MD.                               | <b>cfld md</b> <i>md-name</i> <b>level</b> <i>level-value</i>                                     | Not created by default. |
| 3. Create an MA.                               | <b>cfld ma</b> <i>ma-name</i> <b>md</b> <i>md-name</i> <b>vlan</b> <i>vlan-id</i>                 | Not created by default. |
| 4. Create a service instance with the MD name. | <b>cfld service-instance</b> <i>instance-id</i> <b>md</b> <i>md-name</i> <b>ma</b> <i>ma-name</i> | Not created by default. |

### Creating a service instance without the MD name

When you create a service instance without the MD name, the system automatically creates the MA and MD for the service instance.

To create a service instance without the MD name:

| Step                                              | Command                                                                                                                                                                                                   | Remarks                 |
|---------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------|
| 1. Enter system view.                             | <b>system-view</b>                                                                                                                                                                                        | N/A                     |
| 2. Create a service instance without the MD name. | <b>cfld service-instance</b> <i>instance-id</i><br><b>maid format</b> { <b>icc-based</b> <i>ma-name</i><br>  <b>string</b> <i>ma-name</i> } <b>level</b> <i>level-value</i><br><b>vlan</b> <i>vlan-id</i> | Not created by default. |

## Configuring MEPs

CFD is implemented through various operations on MEPs. As a MEP is configured on a service instance, the MD level and VLAN attribute of the service instance become the attribute of the MEP.

Before creating MEPs, configure the MEP list. An MEP list is a collection of local MEPs allowed to be configured in an MA and the remote MEPs to be monitored.

**IMPORTANT:**

You cannot create a MEP if the MEP ID is not included in the MEP list of the service instance.

To configure a MEP:

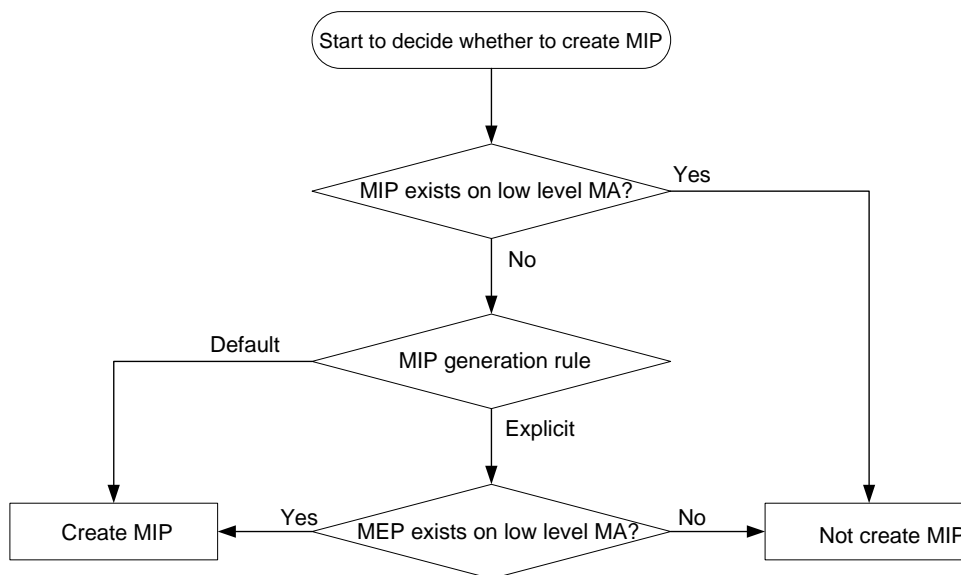
| Step                                      | Command                                                                                                         | Remarks                                |
|-------------------------------------------|-----------------------------------------------------------------------------------------------------------------|----------------------------------------|
| 1. Enter system view.                     | <b>system-view</b>                                                                                              | N/A                                    |
| 2. Configure a MEP list.                  | <b>bfd meplist</b> <i>mep-list</i><br><b>service-instance</b> <i>instance-id</i>                                | By default, no MEP list is configured. |
| 3. Enter Layer 2 Ethernet interface view. | <b>interface</b> <i>interface-type</i><br><i>interface-number</i>                                               | N/A                                    |
| 4. Create a MEP.                          | <b>bfd mep</b> <i>mep-id</i> <b>service-instance</b><br><i>instance-id</i> { <b>inbound</b>   <b>outbound</b> } | Not configured by default.             |
| 5. Enable the MEP.                        | <b>bfd mep</b> <b>service-instance</b><br><i>instance-id</i> <b>mep</b> <i>mep-id</i> <b>enable</b>             | Disabled by default.                   |

## Configuring MIP generation rules

As functional entities in a service instance, MIPs respond to various CFD frames, such as LTM frames, LBM frames, 1DM frames, DMM frames, and TST frames.

MIPs are generated on each port automatically according to related MIP generation rules. If a port has no MIP, the system will check the MAs in each MD (from low to high levels) and follow the procedure described in Figure 7 to create or not to create MIPs (within the same VLAN):

**Figure 7 Procedure of creating MIPs**



You can choose appropriate MIP generation rules based on your network design.

To configure the rules for generating MIPs:

| Step                                        | Command                                                                                                | Remarks                                                                    |
|---------------------------------------------|--------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------|
| 1. Enter system view.                       | <b>system-view</b>                                                                                     | N/A                                                                        |
| 2. Configure the rules for generating MIPs. | <b>cfd mip-rule</b> { <b>explicit</b>   <b>default</b> }<br><b>service-instance</b> <i>instance-id</i> | By default, neither MIPs nor the rules for generating MIPs are configured. |

Any of the following actions or cases can cause MIPs to be created or deleted after you have configured the **cfd mip-rule** command:

- Enabling CFD (use the **cfd enable** command)
- Creating or deleting the MEPs on a port
- Changes occur to the VLAN attribute of a port
- The rule specified in the **cfd mip-rule** command changes

## Configuring CFD functions

### Configuration prerequisites

Complete basic CFD settings.

### Configuring CC on MEPs

After the CC function is configured, MEPs can send CCM frames to one another to check the connectivity between them.

You must configure CC before configuring other CFD functions.

#### CAUTION:

On different devices, the MEPs belonging to the same MD and MA should be configured with the same CCM transmission interval.

To configure CC on a MEP:

| Step                                                                    | Command                                                                                     | Remarks                                                 |
|-------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|---------------------------------------------------------|
| 1. Enter system view.                                                   | <b>system-view</b>                                                                          | N/A                                                     |
| 2. Configure the interval field value in the CCM messages sent by MEPs. | <b>cfd cc interval</b> <i>interval-value</i><br><b>service-instance</b> <i>instance-id</i>  | Optional.<br>By default, the interval field value is 4. |
| 3. Enter Layer 2 Ethernet interface view.                               | <b>interface</b> <i>interface-type</i><br><i>interface-number</i>                           | N/A                                                     |
| 4. Enable CCM sending on a MEP.                                         | <b>cfd cc service-instance</b> <i>instance-id</i><br><b>mep</b> <i>mep-id</i> <b>enable</b> | Disabled by default.                                    |

**Table 9 Relationship between the interval field value, the interval between CCM messages, and the timeout time of the remote MEP**

| The interval field value | The interval between CCM messages | The timeout time of the remote MEP |
|--------------------------|-----------------------------------|------------------------------------|
| 4                        | 1 second                          | 3.5 seconds                        |
| 5                        | 10 second                         | 35 seconds                         |
| 6                        | 60 seconds                        | 210 seconds                        |
| 7                        | 600 seconds                       | 2100 seconds                       |

## Configuring LB on MEPs

The LB function can verify the link state between the local MEP and the remote MEP or MIP.

To configure LB on a MEP:

| Task       | Command                                                                                                                                                                                       | Remarks                                                                                                                                                                   |
|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Enable LB. | <b>cfd loopback service-instance</b><br><i>instance-id mep mep-id</i> { <b>target-mep</b><br><i>target-mep-id</i>   <b>target-mac</b><br><i>mac-address</i> } [ <b>number</b> <i>number</i> ] | Disabled by default.<br>Available in any view<br>The <b>target-mep</b> <i>target-mep-id</i> option is not supported if an outward-facing MEP is configured on the switch. |

## Configuring LT on MEPs

LT can trace the path between the source and target MEPs and can also locate link faults by sending LT messages automatically. The two functions are implemented in the following way:

- To implement the first function, the source MEP first sends LTM messages to the target MEP. Based on the LTR messages in response to the LTM messages, the path between the two MEPs can be identified.
- In the latter case, after LT messages automatic sending is enabled, if the source MEP fails to receive the CCM frames from the target MEP within 3.5 times the transmission interval, the link between the two is considered faulty, and LTM frames (with the target MEP as the destination and the TTL field in the LTM frames set to the maximum value 255) will be sent out. Based on the LTRs that the MIPs return, the fault source can be located.

To configure LT on MEPs:

| Step                                                    | Command                                                                                                                                                                                                              | Remarks                                                                                                                                           |
|---------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Find the path between a source MEP and a target MEP. | <b>cfd linktrace service-instance</b><br><i>instance-id mep mep-id</i><br>{ <b>target-mep</b> <i>target-mep-id</i>  <br><b>target-mac</b> <i>mac-address</i> } [ <b>tth</b><br><i>tth-value</i> ] [ <b>hw-only</b> ] | Available in any view<br>The <b>target-mep</b> <i>target-mep-id</i> option is not supported if an outward-facing MEP is configured on the switch. |
| 2. Enter system view.                                   | <b>system-view</b>                                                                                                                                                                                                   | N/A                                                                                                                                               |
| 3. Enable LT messages automatic sending.                | <b>cfd linktrace auto-detection</b> [ <b>size</b><br><i>size-value</i> ]                                                                                                                                             | Disabled by default.                                                                                                                              |



## Configuring AIS

The AIS function suppresses the number of error alarms reported by MEPs.

To make an MEP in the service instance send AIS frames, you must configure the AIS frame transmission level to be higher than the MD level of the MEP.

Enable AIS and configure the proper AIS frame transmission level on the target MEP, so the target MEP can suppress the error alarms and send the AIS frame to the MD of a higher level. If you enable AIS but do not configure the proper AIS frame transmission level on the target MEP, the target MEP can suppress the error alarms, but cannot send the AIS frames.

To configure AIS:

| Step                                              | Command                                                                              | Remarks                           |
|---------------------------------------------------|--------------------------------------------------------------------------------------|-----------------------------------|
| 1. Enter system view.                             | <b>system-view</b>                                                                   | N/A                               |
| 2. Enable AIS.                                    | <b>cfd ais enable</b>                                                                | Disabled by default.              |
| 3. Configure the AIS frame transmission level.    | <b>cfd ais level</b> <i>level-value</i> <b>service-instance</b> <i>instance-id</i>   | Not configured by default.        |
| 4. Configure the AIS frame transmission interval. | <b>cfd ais period</b> <i>period-value</i> <b>service-instance</b> <i>instance-id</i> | Optional.<br>1 second by default. |

## Configuring LM

The LM function measures frame loss between MEPs, including the number of lost frames, the frame loss ratio, and the average number of lost frames for the source and target MEPs.

To configure LM:

| Step                  | Command                                                                                                                                                                                          | Remarks                                                                                                                                          |
|-----------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view. | <b>system-view</b>                                                                                                                                                                               | N/A                                                                                                                                              |
| 2. Configure LM.      | <b>cfd slm service-instance</b> <i>instance-id</i><br><b>mep</b> <i>mep-id</i> { <b>target-mac</b> <i>mac-address</i>   <b>target-mep</b> <i>target-mep-id</i> } [ <b>number</b> <i>number</i> ] | Disabled by default.<br>The <b>target-mep</b> <i>target-mep-id</i> option is not supported if an outward-facing MEP is configured on the switch. |

### NOTE:

The LM function takes effect only in CFD IEEE 802.1ag.

## Configuring one-way DM

The one-way DM function measures the one-way frame delay between two MEPs, and monitors and manages the link transmission performance.

The one-way DM function takes effect only in CFD IEEE 802.1ag.

One-way DM requires that the clocks at the transmitting MEP and the receiving MEP be synchronized. For the purpose of frame delay variation measurement, the requirement for clock synchronization can be relaxed.

To view the test result, use the **display cfd dm one-way history** command on the target MEP.

To configure one-way DM:

| Step                     | Command                                                                                                                                                                                         | Remarks                                                                                                                                          |
|--------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.    | <b>system-view</b>                                                                                                                                                                              | N/A                                                                                                                                              |
| 2. Configure one-way DM. | <b>cfd dm one-way service-instance</b><br><i>instance-id mep mep-id</i> { <b>target-mac</b><br><i>mac-address</i>   <b>target-mep</b><br><i>target-mep-id</i> } [ <b>number</b> <i>number</i> ] | Disabled by default.<br>The <b>target-mep</b> <i>target-mep-id</i> option is not supported if an outward-facing MEP is configured on the switch. |

## Configuring two-way DM

The two-way DM function measures the two-way frame delay, average two-way frame delay, and two-way frame delay variation between two MEPs, and monitors and manages the link transmission performance.

To configure two-way DM:

| Step                     | Command                                                                                                                                                                                         | Remarks                                                                                                                                          |
|--------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.    | <b>system-view</b>                                                                                                                                                                              | N/A                                                                                                                                              |
| 2. Configure two-way DM. | <b>cfd dm two-way service-instance</b><br><i>instance-id mep mep-id</i> { <b>target-mac</b><br><i>mac-address</i>   <b>target-mep</b><br><i>target-mep-id</i> } [ <b>number</b> <i>number</i> ] | Disabled by default.<br>The <b>target-mep</b> <i>target-mep-id</i> option is not supported if an outward-facing MEP is configured on the switch. |

### NOTE:

The two-way DM function is available only under the IEEE 802.1ag standard version of CFD.

## Configuring TST

The TST function detects bit errors on a link, and monitors and manages the link transmission performance.

To configure TST:

| Step                  | Command                                                                                                                                                                                                                                                                                                                                  | Remarks                                                                                                                                          |
|-----------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view. | <b>system-view</b>                                                                                                                                                                                                                                                                                                                       | N/A                                                                                                                                              |
| 2. Configure TST.     | <b>cfd tst service-instance</b> <i>instance-id</i><br><b>mep</b> <i>mep-id</i> { <b>target-mac</b><br><i>mac-address</i>   <b>target-mep</b><br><i>target-mep-id</i> } [ <b>number</b> <i>number</i> ]<br>[ <b>length-of-test</b> <i>length</i> ]<br>[ <b>pattern-of-test</b> { <b>all-zero</b>   <b>prbs</b> }<br>[ <b>with-crc</b> ] ] | Disabled by default.<br>The <b>target-mep</b> <i>target-mep-id</i> option is not supported if an outward-facing MEP is configured on the switch. |

**NOTE:**

- The TST function takes effect only in CFD IEEE 802.1ag.
- To view the test result, use the **display cfd tst** command on the target MEP.

## Displaying and maintaining CFD

| Task                                                                                         | Command                                                                                                                                                                                              | Remarks               |
|----------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Display CFD and AIS status.                                                                  | <b>display cfd status</b> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                                                                                         | Available in any view |
| Display the CFD protocol version.                                                            | <b>display cfd version</b> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                                                                                        | Available in any view |
| Display MD configuration information.                                                        | <b>display cfd md</b> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                                                                                             | Available in any view |
| Display MA configuration information.                                                        | <b>display cfd ma</b> [ [ <i>ma-name</i> ] <b>md</b> { <i>md-name</i>   <i>level level-value</i> } ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]              | Available in any view |
| Display service instance configuration information.                                          | <b>display cfd service-instance</b> [ <i>instance-id</i> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                                                        | Available in any view |
| Display MEP list in a service instance.                                                      | <b>display cfd meplist</b> [ <b>service-instance</b> <i>instance-id</i> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                                         | Available in any view |
| Display MP information.                                                                      | <b>display cfd mp</b> [ <b>interface</b> <i>interface-type</i> <i>interface-number</i> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                          | Available in any view |
| Display the attribute and running information of the MEPs.                                   | <b>display cfd mep</b> <i>mep-id</i> <b>service-instance</b> <i>instance-id</i> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                                   | Available in any view |
| Display LTR information received by a MEP.                                                   | <b>display cfd linktrace-reply</b> [ <b>service-instance</b> <i>instance-id</i> [ <b>mep</b> <i>mep-id</i> ] ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]    | Available in any view |
| Display the information of a remote MEP.                                                     | <b>display cfd remote-mep</b> <b>service-instance</b> <i>instance-id</i> <b>mep</b> <i>mep-id</i> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                 | Available in any view |
| Display the content of the LTR messages received as responses to the automatically sent LTM. | <b>display cfd linktrace-reply auto-detection</b> [ <b>size</b> <i>size-value</i> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                               | Available in any view |
| Display the AIS configuration and information on the specified MEP.                          | <b>display cfd ais</b> [ <b>service-instance</b> <i>instance-id</i> [ <b>mep</b> <i>mep-id</i> ] ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]                | Available in any view |
| Display the one-way DM result on the specified MEP.                                          | <b>display cfd dm one-way history</b> [ <b>service-instance</b> <i>instance-id</i> [ <b>mep</b> <i>mep-id</i> ] ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] | Available in any view |

| Task                                              | Command                                                                                                                                                                               | Remarks                |
|---------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|
| Display the TST result on the specified MEP.      | <b>display cfd tst</b> [ <b>service-instance</b> <i>instance-id</i> [ <b>mep</b> <i>mep-id</i> ] ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] | Available in any view  |
| Clear the one-way DM result on the specified MEP. | <b>reset cfd dm one-way history</b> [ <b>service-instance</b> <i>instance-id</i> [ <b>mep</b> <i>mep-id</i> ] ]                                                                       | Available in user view |
| Clear the TST result on the specified MEP.        | <b>reset cfd tst</b> [ <b>service-instance</b> <i>instance-id</i> [ <b>mep</b> <i>mep-id</i> ] ]                                                                                      | Available in user view |

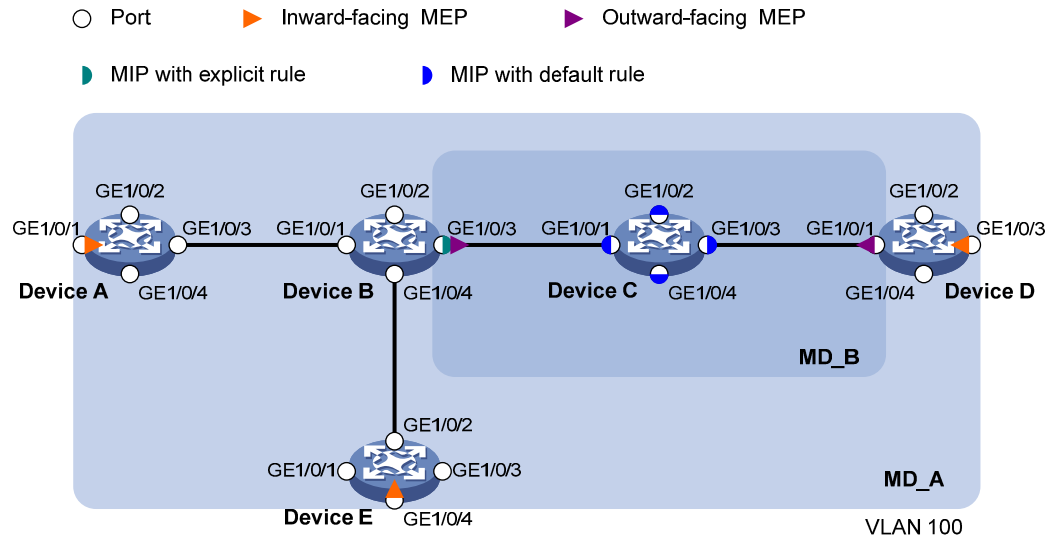
## CFD configuration example

### Network requirements

As shown in [Figure 8](#):

- The network comprises five devices and is divided into two MDs: MD\_A (level 5) and MD\_B (level 3). All ports belong to VLAN 100, and the MAs in the two MDs all serve VLAN 100. Suppose the MAC addresses of Device A through Device E are 0010-FC00-6511, 0010-FC00-6512, 0010-FC00-6513, 0010-FC00-6514, and 0010-FC00-6515.
- MD\_A has three edge ports: GigabitEthernet 1/0/1 on Device A, GigabitEthernet 1/0/3 on Device D, and GigabitEthernet 1/0/4 on Device E, and they are all inward-facing MEPs. MD\_B has two edge ports: GigabitEthernet 1/0/3 on Device B and GigabitEthernet 1/0/1 on Device D, and they are both outward-facing MEPs.
- In MD\_A, Device B is designed to have MIPs when its port is configured with low level MEPs. Port GigabitEthernet 1/0/3 is configured with MEPs of MD\_B, and the MIPs of MD\_A can be configured on this port. You should configure the MIP generation rule of MD\_A as explicit.
- The MIPs of MD\_B are designed on Device C, and are configured on all ports. You should configure the MIP generation rule as default.
- Configure CC to monitor the connectivity among all the MEPs in MD\_A and MD\_B. Configure to use LB to locate link faults, and use the AIS function to suppress the error alarms reported.
- After the status information of the entire network is obtained, use LT, LM, one-way DM, two-way DM, and TST to detect link faults.

**Figure 8 Network diagram**



### Configuration procedure

1. Configure a VLAN and assign ports to it:

On each device shown in [Figure 8](#), create VLAN 100, and assign ports GigabitEthernet 1/0/1 through GigabitEthernet 1/0/4 to VLAN 100.

2. Enable CFD:

# Enable CFD on Device A.

```
<DeviceA> system-view
[DeviceA] cfd enable
```

Enable CFD on Device B through Device E using the same method.

3. Configure service instances:

# Create MD\_A (level 5) on Device A, create MA\_A, which serves VLAN 100, in MD\_A, and create service instance 1 for MD\_A and MA\_A.

```
[DeviceA] cfd md MD_A level 5
[DeviceA] cfd ma MA_A md MD_A vlan 100
[DeviceA] cfd service-instance 1 md MD_A ma MA_A
```

Configure Device E as you configure Device A.

# Create MD\_A (level 5) on Device B, create MA\_A, which serves VLAN 100, in MD\_A, and then create service instance 1 for MD\_A and MA\_A. In addition, create MD\_B (level 3), create MA\_B, which serves VLAN 100, in MD\_B, and then create service instance 2 for MD\_B and MA\_B.

```
[DeviceB] cfd md MD_A level 5
[DeviceB] cfd ma MA_A md MD_A vlan 100
[DeviceB] cfd service-instance 1 md MD_A ma MA_A
[DeviceB] cfd md MD_B level 3
[DeviceB] cfd ma MA_B md MD_B vlan 100
[DeviceB] cfd service-instance 2 md MD_B ma MA_B
```

Configure Device D as you configure Device B.

# Create MD\_B (level 3) on Device C, create MA\_B, which serves VLAN 100, in MD\_B, and then create service instance 2 for MD\_B and MA\_B.

```
[DeviceC] cfd md MD_B level 3
```

```
[DeviceC] cfd ma MA_B md MD_B vlan 100
[DeviceC] cfd service-instance 2 md MD_B ma MA_B
```

#### 4. Configure MEPs:

# On Device A, configure a MEP list in service instance 1. Create and enable inward-facing MEP 1001 in service instance 1 on GigabitEthernet 1/0/1.

```
[DeviceA] cfd meplist 1001 4002 5001 service-instance 1
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] cfd mep 1001 service-instance 1 inbound
[DeviceA-GigabitEthernet1/0/1] cfd mep service-instance 1 mep 1001 enable
[DeviceA-GigabitEthernet1/0/1] quit
```

# On Device B, configure a MEP list in service instances 1 and 2, respectively. Create and enable outward-facing MEP 2001 in service instance 2 on GigabitEthernet 1/0/3.

```
[DeviceB] cfd meplist 1001 4002 5001 service-instance 1
[DeviceB] cfd meplist 2001 4001 service-instance 2
[DeviceB] interface gigabitethernet 1/0/3
[DeviceB-GigabitEthernet1/0/3] cfd mep 2001 service-instance 2 outbound
[DeviceB-GigabitEthernet1/0/3] cfd mep service-instance 2 mep 2001 enable
[DeviceB-GigabitEthernet1/0/3] quit
```

# On Device D, configure a MEP list in service instances 1 and 2, respectively. Create and enable outward-facing MEP 4001 in service instance 2 on GigabitEthernet 1/0/1, and then create and enable inward-facing MEP 4002 in service instance 1 on GigabitEthernet 1/0/3.

```
[DeviceD] cfd meplist 1001 4002 5001 service-instance 1
[DeviceD] cfd meplist 2001 4001 service-instance 2
[DeviceD] interface gigabitethernet 1/0/1
[DeviceD-GigabitEthernet1/0/1] cfd mep 4001 service-instance 2 outbound
[DeviceD-GigabitEthernet1/0/1] cfd mep service-instance 2 mep 4001 enable
[DeviceD-GigabitEthernet1/0/1] quit
[DeviceD] interface gigabitethernet 1/0/3
[DeviceD-GigabitEthernet1/0/3] cfd mep 4002 service-instance 1 inbound
[DeviceD-GigabitEthernet1/0/3] cfd mep service-instance 1 mep 4002 enable
[DeviceD-GigabitEthernet1/0/3] quit
```

# On Device E, configure a MEP list in service instance 1. Create and enable inward-facing MEP 5001 in service instance 1 on GigabitEthernet 1/0/4.

```
[DeviceE] cfd meplist 1001 4002 5001 service-instance 1
[DeviceE] interface gigabitethernet 1/0/4
[DeviceE-GigabitEthernet1/0/4] cfd mep 5001 service-instance 1 inbound
[DeviceE-GigabitEthernet1/0/4] cfd mep service-instance 1 mep 5001 enable
[DeviceE-GigabitEthernet1/0/4] quit
```

#### 5. Configure MIPs:

# Configure the MIP generation rule in service instance 1 on Device B as explicit.

```
[DeviceB] cfd mip-rule explicit service-instance 1
```

# Configure the MIP generation rule in service instance 2 on Device C as default.

```
[DeviceC] cfd mip-rule default service-instance 2
```

#### 6. Configure CC:

# On Device A, enable the sending of CCM frames for MEP 1001 in service instance 1 on GigabitEthernet 1/0/1.

```
[DeviceA] interface gigabitethernet 1/0/1
```

```
[DeviceA-GigabitEthernet1/0/1] cfd cc service-instance 1 mep 1001 enable
[DeviceA-GigabitEthernet1/0/1] quit
```

# On Device B, enable the sending of CCM frames for MEP 2001 in service instance 2 on GigabitEthernet 1/0/3.

```
[DeviceB] interface gigabitethernet 1/0/3
[DeviceB-GigabitEthernet1/0/3] cfd cc service-instance 2 mep 2001 enable
[DeviceB-GigabitEthernet1/0/3] quit
```

# On Device D, enable the sending of CCM frames for MEP 4001 in service instance 2 on GigabitEthernet 1/0/1, and enable the sending of CCM frames for MEP 4002 in service instance 1 on GigabitEthernet 1/0/3.

```
[DeviceD] interface gigabitethernet 1/0/1
[DeviceD-GigabitEthernet1/0/1] cfd cc service-instance 2 mep 4001 enable
[DeviceD-GigabitEthernet1/0/1] quit
[DeviceD] interface gigabitethernet 1/0/3
[DeviceD-GigabitEthernet1/0/3] cfd cc service-instance 1 mep 4002 enable
[DeviceD-GigabitEthernet1/0/3] quit
```

# On Device E, enable the sending of CCM frames for MEP 5001 in service instance 1 on GigabitEthernet 1/0/4.

```
[DeviceE] interface gigabitethernet 1/0/4
[DeviceE-GigabitEthernet1/0/4] cfd cc service-instance 1 mep 5001 enable
[DeviceE-GigabitEthernet1/0/4] quit
```

## 7. Configure AIS:

# Enable AIS on Device B, and configure the AIS frame transmission level as 2 and AIS frame transmission interval as 1 second in service instance 2.

```
[DeviceB] cfd ais enable
[DeviceB] cfd ais level 5 service-instance 2
[DeviceB] cfd ais period 1 service-instance 2
```

## Verifying the configuration

### 1. Verify the LB function:

When the CC function detects a link fault, use the LB function to locate the fault.

# Enable LB on Device A to check the status of the link between MEP 1001 and MEP 5001 in service instance 1.

```
[DeviceA] cfd loopback service-instance 1 mep 1001 target-mep 5001
Loopback to 0010-FC00-6515 with the sequence number start from 1001-43404:
Reply from 0010-FC00-6515: sequence number=1001-43404 time=5ms
Reply from 0010-FC00-6515: sequence number=1001-43405 time=5ms
Reply from 0010-FC00-6515: sequence number=1001-43406 time=5ms
Reply from 0010-FC00-6515: sequence number=1001-43407 time=5ms
Reply from 0010-FC00-6515: sequence number=1001-43408 time=5ms
Send:5 Received:5 Lost:0
```

After the whole network status is obtained with the CC function, use the LT function to identify the paths between source and target MEPs or locate faults.

### 2. Verify the LT function:

# Identify the path between MEP 1001 and MEP 5001 in service instance 1 on Device A.

```
[DeviceA] cfd linktrace service-instance 1 mep 1001 target-mep 5001
Linktrace to MEP 5001 with the sequence number 1001-43462
```

| MAC Address    | TTL | Last MAC       | Relay Action |
|----------------|-----|----------------|--------------|
| 0010-FC00-6515 | 63  | 0010-FC00-6512 | Hit          |

### 3. Verify the LM function:

After the CC function obtains the status information of the entire network, use the LM function to test the link status. For example:

# Test the frame loss from MEP 1001 to MEP 4002 in service instance 1 on Device A.

```
[DeviceA] cfd slm service-instance 1 mep 1001 target-mep 4002
Reply from 0010-FC00-6514
Far-end frame loss: 10 Near-end frame loss: 20
Reply from 0010-FC00-6514
Far-end frame loss: 40 Near-end frame loss: 40
Reply from 0010-FC00-6514
Far-end frame loss: 0 Near-end frame loss: 10
Reply from 0010-FC00-6514
Far-end frame loss: 30 Near-end frame loss: 30
```

Average

```
Far-end frame loss: 20 Near-end frame loss: 25
Far-end frame loss rate: 25% Near-end frame loss rate: 32%
Send LMMs: 5 Received: 5 Lost: 0
```

### 4. Verify the one-way DM function:

After the CC function obtains the status information of the entire network, use the one-way DM function to test the one-way frame delay of a link. For example:

# Test the one-way frame delay from MEP 1001 to MEP 4002 in service instance 1 on Device A.

```
[DeviceA] cfd dm one-way service-instance 1 mep 1001 target-mep 4002
Info: 5 1DM frames process is done, please check the result on the remote device.
```

# Display the one-way DM result on MEP 4002 in service instance 1 on Device D.

```
[DeviceD] display cfd dm one-way history service-instance 1 mep 4002
Service instance: 1
MEP ID: 4002
Send 1DM total number: 0
Received 1DM total number: 5
Frame delay: 10ms 9ms 11ms 5ms 5ms
Delay average: 8ms
Delay variation: 5ms 4ms 6ms 0ms 0ms
Variation average: 3ms
```

### 5. Verify the two-way DM function:

After the CC function obtains the status information of the entire network, use the two-way DM function to test the two-way frame delay of a link. For example:

# Test the two-way frame delay from MEP 1001 to MEP 4002 in service instance 1 on Device A.

```
[DeviceA] cfd dm two-way service-instance 1 mep 1001 target-mep 4002
Frame delay:
Reply from 0010-FC00-6514: 10ms
Reply from 0010-FC00-6514: 9ms
Reply from 0010-FC00-6514: 11ms
Reply from 0010-FC00-6514: 5ms
```



Reply from 0010-FC00-6514: 5ms

Average: 8ms

Send DMMs: 5            Received: 5            Lost: 0

Frame delay variation: 5ms 4ms 6ms 0ms 0ms

Average: 3ms

**6. Verify the TST function:**

After the CC function obtains the status information of the entire network, use the TST function to test the bit errors of a link. For example:

# Test the bit errors on the link from MEP 1001 to MEP 4002 in service instance 1 on Device A.

```
[DeviceA] cfd tst service-instance 1 mep 1001 target-mep 4002
```

Info: TST process is done. Please check the result on the remote device.

# Display the TST result on MEP 4002 in service instance 1 on Device D.

```
[DeviceD] display cfd tst service-instance 1 mep 4002
```

Service instance: 1

MEP ID: 4002

Send TST total number: 0

Received TST total number: 5

Received from 0010-FC00-6511, sequence number 1: Bit True

Received from 0010-FC00-6511, sequence number 2: Bit True

Received from 0010-FC00-6511, sequence number 3: Bit True

Received from 0010-FC00-6511, sequence number 4: Bit True

Received from 0010-FC00-6511, sequence number 5: Bit True

# Configuring DLDAP

## DLDP overview

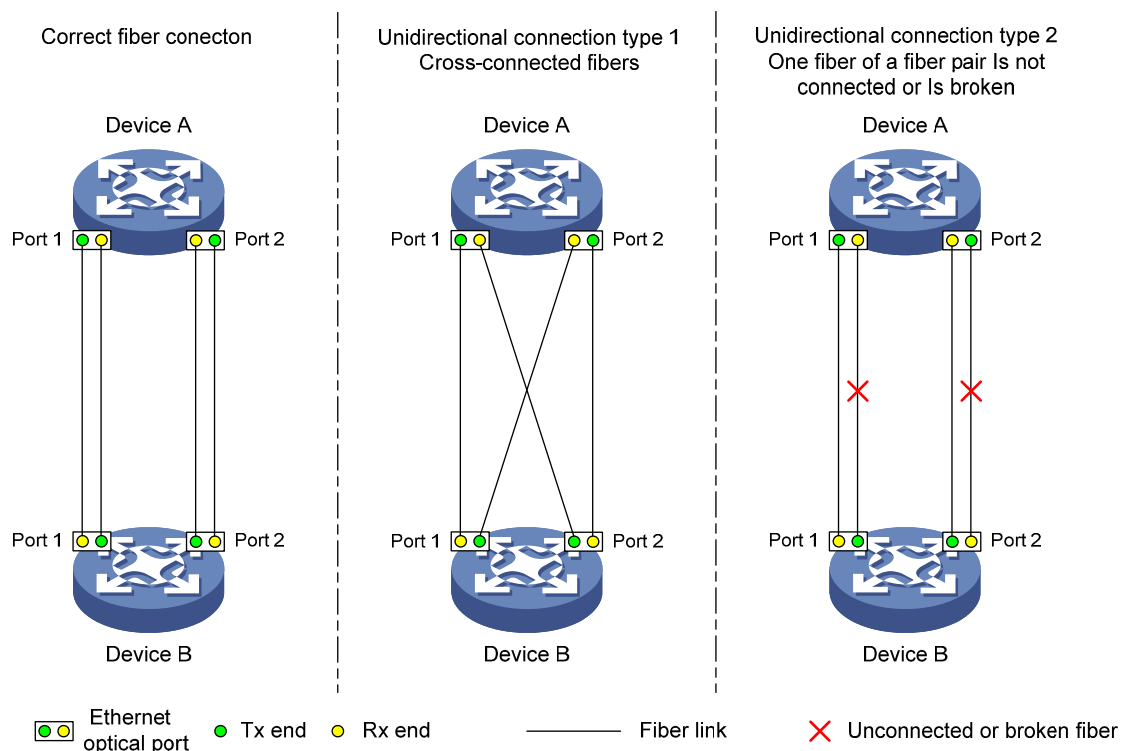
### Background

Unidirectional links occur when one end of a link can receive packets from the other end, but the other end cannot receive packets sent by the first end. Unidirectional links result in problems such as loops in an STP-enabled network.

For example, the link between two switches, Switch A and Switch B, is a bidirectional link when they are connected via a fiber pair, with one fiber used for sending packets from A to B and the other for sending packets from B to A. This link is a two-way link. If one of the fibers gets broken, the link becomes a unidirectional link (one-way link).

There are two types of unidirectional fiber links. One occurs when fibers are cross-connected. The other occurs when a fiber is not connected at one end, or when one fiber of a fiber pair gets broken. [Figure 9](#) shows a correct fiber connection and the two types of unidirectional fiber connection.

**Figure 9 Correct and incorrect fiber connections**



The Device link detection protocol (DLDP) detects unidirectional links (fiber links or twisted-pair links) and can be configured to shut down the related port automatically or prompt users to take actions to avoid network problems.

As a data link layer protocol, DLDP cooperates with physical layer protocols to monitor link status. When the auto-negotiation mechanism provided by the physical layer detects physical signals and faults, DLDP

performs operations such as identifying peer devices, detecting unidirectional links, and shutting down unreachable ports. The auto-negotiation mechanism and DLDAP work together to make sure that physical/logical unidirectional links are detected and shut down, and to prevent failure of other protocols such as STP. If both ends of a link are operating normally at the physical layer, DLDAP detects whether the link is correctly connected at the link layer and whether the two ends can exchange packets properly. This is beyond the capability of the auto-negotiation mechanism at the physical layer.

## How DLDAP works

### DLDAP link states

A device is in one of these DLDAP link states: Initial, Inactive, Active, Advertisement, Probe, Disable, and DelayDown, as described in [Table 10](#).

**Table 10 DLDAP link states**

| State         | Indicates...                                                                                                                                                                                                                                                                             |
|---------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Initial       | DLDAP is disabled.                                                                                                                                                                                                                                                                       |
| Inactive      | DLDAP is enabled, and the link is down.                                                                                                                                                                                                                                                  |
| Active        | DLDAP is enabled and the link is up, or the neighbor entries have been cleared.                                                                                                                                                                                                          |
| Advertisement | All neighbors are bi-directionally reachable or DLDAP has been in active state for more than five seconds. This is a relatively stable state where no unidirectional link has been detected.                                                                                             |
| Probe         | DLDAP enters this state if it receives a packet from an unknown neighbor. In this state, DLDAP sends packets to check whether the link is unidirectional. As soon as DLDAP transits to this state, a probe timer starts and an echo timeout timer starts for each neighbor to be probed. |
| Disable       | A port enters this state when: <ul style="list-style-type: none"> <li>• A unidirectional link is detected.</li> <li>• The contact with the neighbor in enhanced mode gets lost.</li> <li>• In this state, the port does not receive or send packets other than DLDAPDUs.</li> </ul>      |
| DelayDown     | A port in the Active, Advertisement, or Probe DLDAP link state transits to this state rather than removes the corresponding neighbor entry and transits to the Inactive state when it detects a port-down event. When a port transits to this state, the DelayDown timer is triggered.   |

### DLDAP timers

**Table 11 DLDAP timers**

| DLDAP timer         | Description                                                                                                                                                                                                                                                                                                       |
|---------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Active timer        | Determines the interval for sending Advertisement packets with RSY tags, which defaults to 1 second. By default, a device in the active DLDAP link state sends one Advertisement packet with RSY tags every second. The maximum number of advertisement packets with RSY tags that can be sent successively is 5. |
| Advertisement timer | Determines the interval for sending common advertisement packets, which defaults to 5 seconds.                                                                                                                                                                                                                    |
| Probe timer         | Determines the interval for sending Probe packets, which defaults to 1 second. By default, a device in the probe state sends one Probe packet every second. The maximum number of Probe packets that can be sent successively is 10.                                                                              |

| DLDP timer         | Description                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|--------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Echo timer         | <p>This timer is set to 10 seconds. It is triggered when a device transits to the Probe state or when an enhanced detect is launched. When the Echo timer expires and no Echo packet has been received from a neighbor device, the state of the link is set to unidirectional and the device transits to the Disable state. In this case, the device does the following:</p> <p>Sends Disable packets.</p> <p>Either prompts the user to shut down the port or shuts down the port automatically (depending on the DLDP down mode configured).</p> <p>Removes the corresponding neighbor entries.</p>                                                        |
| Entry timer        | <p>When a new neighbor joins, a neighbor entry is created and the corresponding entry timer is triggered. When a DLDP packet is received, the device updates the corresponding neighbor entry and the entry timer.</p> <p>In normal mode, if no packet is received from a neighbor when the corresponding entry timer expires, DLDP sends advertisement packets with RSY tags and removes the neighbor entry.</p> <p>In enhanced mode, if no packet is received from a neighbor when the Entry timer expires, DLDP triggers the enhanced timer.</p> <p>The setting of an Entry timer is three times that of the Advertisement timer.</p>                     |
| Enhanced timer     | <p>In enhanced mode, this timer is triggered if no packet is received from a neighbor when the entry timer expires. Enhanced timer is set to 1 second.</p> <p>After the Enhanced timer is triggered, the device sends up to eight probe packets to the neighbor at a frequency of one packet per second.</p>                                                                                                                                                                                                                                                                                                                                                 |
| DelayDown timer    | <p>A device in Active, Advertisement, or Probe DLDP link state transits to DelayDown state rather than removes the corresponding neighbor entry and transits to the Inactive state when it detects a port-down event.</p> <p>When a device transits to this state, the DelayDown timer is triggered. A device in DelayDown state only responds to port-up events.</p> <p>If a device in the DelayDown state detects a port-up event before the DelayDown timer expires, it resumes its original DLDP state. If not, when the DelayDown timer expires, the device removes the corresponding DLDP neighbor information and transits to the Inactive state.</p> |
| RecoverProbe timer | <p>This timer is set to 2 seconds. A port in the Disable state sends one RecoverProbe packet every two seconds to detect whether a unidirectional link has restored.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

## DLDP mode

DLDP can operate in normal or enhanced mode:

- In normal DLDP mode, when an entry timer expires, the device removes the corresponding neighbor entry and sends an Advertisement packet with the RSY tag.
- In enhanced DLDP mode, when an entry timer expires, the Enhanced timer is triggered and the device tests the neighbor by sending up to eight Probe packets at the frequency of one packet per second. If no Echo packet has been received from the neighbor when the Echo timer expires, the device transits to the Disable state.

Table 12 shows the relationship between the DLDP modes and neighbor entry aging.

**Table 12 DLDAP mode and neighbor entry aging**

| DLDAP mode          | Detecting a neighbor after the corresponding neighbor entry ages out | Removing the neighbor entry immediately after the Entry timer expires | Triggering the Enhanced timer after an Entry timer expires |
|---------------------|----------------------------------------------------------------------|-----------------------------------------------------------------------|------------------------------------------------------------|
| Normal DLDAP mode   | No                                                                   | Yes                                                                   | No                                                         |
| Enhanced DLDAP mode | Yes                                                                  | No                                                                    | Yes                                                        |

Table 13 shows the relationship between DLDAP modes and unidirectional link types.

**Table 13 DLDAP mode and unidirectional link types**

| Unidirectional link type      | Whether it occurs on fibers | Whether it occurs on copper twisted pairs | In which DLDAP mode unidirectional links can be detected                                                                                    |
|-------------------------------|-----------------------------|-------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|
| Cross-connected link          | Yes                         | No                                        | Both normal and enhanced modes                                                                                                              |
| Connectionless or broken link | Yes                         | Yes                                       | Only enhanced mode. The port that can receive signals is in Disable state, and the port that does not receive signals is in Inactive state. |

Enhanced DLDAP mode is designed for addressing black holes. It prevents situations where one end of a link is up and the other is down.

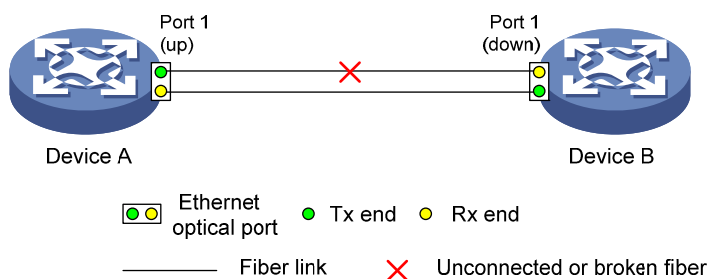
If you configure forced speed and full duplex mode on a port, the situation shown in Figure 10 may occur (take the fiber link for example). Without DLDAP enabled, the port on Device B is actually down but its state cannot be detected by common data link protocols, so the port on Device A is still up. However, in enhanced DLDAP mode, the following occurs:

The port on Device B is in Inactive DLDAP state because it is physically down.

The port on Device A tests the peer port on Device B after the Entry timer for the port on Device B expires.

The port on Device A transits to the Disable state if it does not receive an Echo packet from the port on Device B when the Echo timer expires.

**Figure 10 A scenario for the enhanced DLDAP mode**



### DLDAP authentication mode

You can use DLDAP authentication to prevent network attacks and illegal detection. There are three DLDAP authentication modes.

- Non-authentication:
  - The sending side sets the Authentication field and the Authentication type field of DLDAP packets to 0.
  - The receiving side checks the values of the two fields of received DLDAP packets and drops the packets where the two fields conflict with the corresponding local configuration.
- Simple authentication:
  - Before sending a DLDAP packet, the sending side sets the Authentication field to the user-configured password and sets the Authentication type field to 1.
  - The receiving side checks the values of the two fields in received DLDAP packets and drops any packets where the two fields conflict with the corresponding local configuration.
- MD5 authentication:
  - Before sending a packet, the sending side encrypts the user configured password using MD5 algorithm, assigns the digest to the Authentication field, and sets the Authentication type field to 2.
  - The receiving side checks the values of the two fields in received DLDAP packets and drops any packets where the two fields conflicting with the corresponding local configuration.

## DLDAP processes

1. On a DLDAP-enabled link that is in up state, DLDAP sends DLDAP packets to the peer device and processes the DLDAP packets received from the peer device. DLDAP packets sent vary with DLDAP states.

**Table 14 DLDAP packet types and DLDAP states**

| DLDAP state   | Type of DLDAP packets sent                  |
|---------------|---------------------------------------------|
| Active        | Advertisement packet with RSY tag           |
| Advertisement | Normal Advertisement packet                 |
| Probe         | Probe packet                                |
| Disable       | Disable packet and then RecoverProbe packet |

### NOTE:

A device sends Flush packets when it transits to the Initial state from the Active, Advertisement, Probe, or DelayDown state but does not send them when it transits to the Initial state from Inactive or Disable state.

2. A received DLDAP packet is processed with the following methods:
  - In any of the three authentication modes, the packet is dropped if it fails to pass the authentication.
  - The packet is dropped if the setting of the interval to send Advertisement packets it carries conflicts with the corresponding local setting.
  - Other processes are as shown in [Table 15](#).

**Table 15 Procedures for processing different types of DLDP packets received**

| Packet type                       | Processing procedure                                               |                                                                                                                                                                                                                                                                                                                                                                                                  |
|-----------------------------------|--------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Advertisement packet with RSY tag | Retrieves the neighbor information                                 | If the corresponding neighbor entry does not exist, creates the neighbor entry, triggers the Entry timer, and transits to Probe state.                                                                                                                                                                                                                                                           |
|                                   |                                                                    | If the corresponding neighbor entry already exists, resets the Entry timer and transits to Probe state.                                                                                                                                                                                                                                                                                          |
| Normal Advertisement packet       | Retrieves the neighbor information                                 | If the corresponding neighbor entry does not exist, creates the neighbor entry, triggers the Entry timer, and transits to Probe state.                                                                                                                                                                                                                                                           |
|                                   |                                                                    | If the corresponding neighbor entry already exists, resets the Entry timer.                                                                                                                                                                                                                                                                                                                      |
| Flush packet                      | Determines whether or not the local port is in Disable state       | If yes, performs no processing.                                                                                                                                                                                                                                                                                                                                                                  |
|                                   |                                                                    | If no, removes the corresponding neighbor entry (if any).                                                                                                                                                                                                                                                                                                                                        |
| Probe packet                      | Retrieves the neighbor information                                 | If the corresponding neighbor entry does not exist, creates the neighbor entry, transits to Probe state, and returns Echo packets.                                                                                                                                                                                                                                                               |
|                                   |                                                                    | If the corresponding neighbor entry already exists, resets the Entry timer and returns Echo packets.                                                                                                                                                                                                                                                                                             |
| Echo packet                       | Retrieves the neighbor information                                 | If the corresponding neighbor entry does not exist, creates the neighbor entry, triggers the Entry timer, and transits to Probe state.                                                                                                                                                                                                                                                           |
|                                   |                                                                    | The corresponding neighbor entry already exists                                                                                                                                                                                                                                                                                                                                                  |
| Disable packet                    | Checks whether the local port is in Disable state                  | If yes, performs no processing.                                                                                                                                                                                                                                                                                                                                                                  |
|                                   |                                                                    | If not, sets the state of the corresponding neighbor to unidirectional, and then checks the state of other neighbors. If all the neighbors are unidirectional, transitions the local port to the Disable state. If the state of some neighbors is unknown, waits until the state of these neighbors is determined. If bidirectional neighbors are present, removes all unidirectional neighbors. |
| RecoverProbe packet               | Checks whether the local port is in Disable or Advertisement state | If not, performs no processing.                                                                                                                                                                                                                                                                                                                                                                  |
|                                   |                                                                    | If yes, returns RecoverEcho packets.                                                                                                                                                                                                                                                                                                                                                             |
| RecoverEcho packet                | Checks whether the local port is in Disable state                  | If not, performs no processing.                                                                                                                                                                                                                                                                                                                                                                  |
|                                   |                                                                    | If yes, the local port transits to Active state if the neighbor information the packet carries is consistent with the local port information.                                                                                                                                                                                                                                                    |

| Packet type     | Processing procedure                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|-----------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                 | If not, performs no processing.                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| LinkDown packet | Checks whether the local port operates in Enhanced mode<br>If yes and the local port is not in Disable state, sets the state of the corresponding neighbor to unidirectional, and then checks the state of other neighbors. If all the neighbors are unidirectional, transitions the local port to the Disable state. If the state of some neighbors is unknown, waits until the state of these neighbors is determined. If bidirectional neighbors are present, removes all unidirectional neighbors. |

3. If no echo packet is received from the neighbor, DLDP performs the following processing.

**Table 16 DLDP process when no echo packet is received from the neighbor**

| No echo packet received from the neighbor                                 | Processing procedure                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|---------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| In normal mode, no echo packet is received when the Echo timer expires.   | DLDP sets the state of the corresponding neighbor to unidirectional, and then checks the state of other neighbors: <ul style="list-style-type: none"> <li>If all the neighbors are unidirectional, removes all the neighbors, transitions to the Disable state, outputs log and tracking information, and sends Disable packets. In addition, depending on the user-defined DLDP down mode, shuts down the local port or prompts users to shut down the port.</li> <li>If the state of some neighbors is unknown, waits until the state of these neighbors is determined.</li> <li>If bidirectional neighbors are present, removes all unidirectional neighbors.</li> </ul> |
| In enhanced mode, no echo packet is received when the Echo timer expires. |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |

### Link auto-recovery mechanism

If the port shutdown mode upon detection of a unidirectional link is set to **auto**, DLDP automatically sets the state of the port, where a unidirectional link is detected, to DLDP down. A DLDP down port cannot forward data traffic or send/receive any PDUs except DLDAPDUs.

On a DLDP down port, DLDP monitors the unidirectional link. Once DLDP finds out that the state of the link has restored to bidirectional, it brings up the port. The specific process is:

The DLDP down port sends out a RecoverProbe packet, which carries only information about the local port, every two seconds. Upon receiving the RecoverProbe packet, the remote end returns a RecoverEcho packet. Upon receiving the RecoverEcho packet, the local port checks whether neighbor information in the RecoverEcho packet is the same as the local port information. If they are the same, the link between the local port and the neighbor is considered to have been restored to a bidirectional link, and the port will transit from Disable state to Active state and re-establish relationship with the neighbor.

Only DLDP down ports can send and process Recover packets, including RecoverProbe packets and RecoverEcho packets. If related ports are manually shut down with the **shutdown** command, the auto-recovery mechanism will not take effect.

### DLDP neighbor state

A DLDP neighbor can be in one of the three states described in [Table 17](#).



**Table 17 Description on DLDAP neighbor states**

| <b>DLDAP neighbor state</b> | <b>Description</b>                                                                                                                                                                                                           |
|-----------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Unknown                     | A neighbor is in this state when it is just detected and is being probed. A neighbor is in this state only when it is being probed. It transits to Two way state or Unidirectional state after the probe operation finishes. |
| Two way                     | A neighbor is in this state after it receives response from its peer. This state indicates the link is a two-way link.                                                                                                       |
| Unidirectional              | A neighbor is in this state when the link connecting it is detected to be a unidirectional link. After a device transits to this state, the corresponding neighbor entries maintained on other devices are removed.          |

## DLDAP configuration task list

For DLDAP to work properly, enable DLDAP on both sides and make sure these settings are consistent: the interval to send Advertisement packets, DLDAP authentication mode, and password.

DLDAP does not process any link aggregation control protocol (LACP) events. The links in an aggregation are treated as individual links in DLDAP.

Make sure the DLDAP version running on devices on the two ends are the same.

Complete the following tasks to configure DLDAP:

| <b>Task</b>                                                                    | <b>Remarks</b> |
|--------------------------------------------------------------------------------|----------------|
| <a href="#">Configuring the duplex mode and speed of an Ethernet interface</a> | Required       |
| <a href="#">Enabling DLDAP</a>                                                 | Required       |
| <a href="#">Setting DLDAP mode</a>                                             | Optional       |
| <a href="#">Setting the interval to send advertisement packets</a>             | Optional       |
| <a href="#">Setting the delaydown timer</a>                                    | Optional       |
| <a href="#">Setting the port shutdown mode</a>                                 | Optional       |
| <a href="#">Configuring DLDAP authentication</a>                               | Optional       |
| <a href="#">Resetting DLDAP state</a>                                          | Optional       |

## Configuring the duplex mode and speed of an Ethernet interface

To make sure that DLDAP works properly on a link, you must configure the full duplex mode for the ports at two ends of the link, and configure a speed for the two ports, rather than letting them negotiate a speed.

For more information about the **duplex** and **speed** commands, see *Layer 2—LAN Switching Command Reference*.

# Enabling DLDAP

To properly configure DLDAP on the device, first enable DLDAP globally, and then enable it on each port.

To enable DLDAP:

| Step                                                         | Command                                                                                                                                                                               | Remarks                                                                                                                                                                                        |
|--------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                                        | <b>system-view</b>                                                                                                                                                                    | N/A                                                                                                                                                                                            |
| 2. Enable DLDAP globally.                                    | <b>dldap enable</b>                                                                                                                                                                   | Globally disabled by default.                                                                                                                                                                  |
| 3. Enter Layer 2 Ethernet interface view or port group view. | Enter Layer 2 Ethernet interface view:<br><b>interface</b> <i>interface-type</i> <i>interface-number</i><br>Enter port group view:<br><b>port-group manual</b> <i>port-group-name</i> | Use either approach.<br>Configurations made in Layer 2 Ethernet interface view apply to the current port only.<br>Configurations made in port group view apply to all ports in the port group. |
| 4. Enable DLDAP.                                             | <b>dldap enable</b>                                                                                                                                                                   | Disabled on a port by default.                                                                                                                                                                 |

## NOTE:

- DLDAP takes effect only on Ethernet interfaces (fiber or copper).
- DLDAP can detect unidirectional links only after all physical links are connected. Therefore, before enabling DLDAP, make sure that optical fibers or copper twisted pairs are connected.

# Setting DLDAP mode

DLDAP operates in normal or enhanced mode.

In normal mode, DLDAP does not actively detect neighbors when the corresponding neighbor entries age out.

In enhanced mode, DLDAP actively detects neighbors when the corresponding neighbor entries age out.

To set DLDAP mode:

| Step                  | Command                                     | Remarks                         |
|-----------------------|---------------------------------------------|---------------------------------|
| 1. Enter system view. | <b>system-view</b>                          | N/A                             |
| 2. Set DLDAP mode.    | <b>dldap work-mode { enhance   normal }</b> | Optional.<br>Normal by default. |

# Setting the interval to send advertisement packets

DLDAP detects unidirectional links by sending Advertisement packets. To make sure that DLDAP can detect unidirectional links promptly without affecting network performance, set the advertisement interval appropriately depending on your network environment. The interval should be set shorter than one third of the STP convergence time. If the interval is too long, STP loops may occur before unidirectional links are detected and shut down. If the interval is too short, the number of advertisement packets will increase. HP recommends that you use the default interval in most cases.

To set the interval to send Advertisement packets:

| Step                                               | Command                   | Remarks                            |
|----------------------------------------------------|---------------------------|------------------------------------|
| 1. Enter system view.                              | <b>system-view</b>        | N/A                                |
| 2. Set the interval to send Advertisement packets. | <b>dldp interval time</b> | Optional.<br>5 seconds by default. |

**NOTE:**

- The interval for sending Advertisement packets applies to all DLDAP-enabled ports.
- To enable DLDAP to operate properly, make sure the intervals for sending Advertisement packets on both sides of a link are the same.

## Setting the delaydown timer

On some ports, when the Tx line fails, the port goes down and then comes up again, causing optical signal jitters on the Rx line. When a port goes down due to a Tx failure, the device transits to the DelayDown state instead of the Inactive state to prevent the corresponding neighbor entries from being removed. At the same time, the device triggers the DelayDown timer. If the port goes up before the timer expires, the device restores the original state; if the port remains down when the timer expires, the device transits to the Inactive state.

To set the DelayDown timer:

| Step                        | Command                          | Remarks                           |
|-----------------------------|----------------------------------|-----------------------------------|
| 1. Enter system view.       | <b>system-view</b>               | N/A                               |
| 2. Set the DelayDown timer. | <b>dldp delaydown-timer time</b> | Optional.<br>1 second by default. |

**NOTE:**

DelayDown timer setting applies to all DLDAP-enabled ports.

## Setting the port shutdown mode

On detecting a unidirectional link, the ports can be shut down in one of the following two modes:

- **Manual mode**—This mode applies to low performance networks, where normal links may be treated as unidirectional links. It protects data traffic transmission against false unidirectional links. In this mode, DLDAP only detects unidirectional links but does not automatically shut down unidirectional link ports. Instead, the DLDAP state machine generates log and traps to prompt you to manually shut down unidirectional link ports with the **shutdown** command. HP recommends that you do as prompted. Then the DLDAP state machine transits to the Disable state.
- **Auto mode**—In this mode, when a unidirectional link is detected, DLDAP transits to Disable state, generates log and traps, and sets the port state to DLDAP Down.

On a port with both remote OAM loopback and DLDAP enabled, if the port shutdown mode is auto mode, the port will be shut down by DLDAP when it receives a packet sent by itself, causing remote OAM loopback to operate improperly. To prevent this, set the port shutdown mode to manual mode.

If the device is busy, or the CPU usage is high, normal links may be treated as unidirectional links. In this case, you can set the port shutdown mode to manual mode to alleviate the impact caused by false unidirectional link report.

To set port shutdown mode:

| Step                       | Command                                                             | Remarks                              |
|----------------------------|---------------------------------------------------------------------|--------------------------------------|
| 1. Enter system view.      | <b>system-view</b>                                                  | N/A                                  |
| 2. Set port shutdown mode. | <b>dldp unidirectional-shutdown</b> { <b>auto</b>   <b>manual</b> } | Optional.<br><b>auto</b> by default. |

## Configuring DLDAP authentication

You can guard your network against attacks and malicious probes by configuring an appropriate DLDAP authentication mode, which can be simple authentication or MD5 authentication. If your network is safe, you can choose not to authenticate.

To enable DLDAP to operate properly, make sure that DLDAP authentication modes and passwords on both sides of a link are the same.

To configure DLDAP authentication:

| Step                               | Command                                                                                           | Remarks                 |
|------------------------------------|---------------------------------------------------------------------------------------------------|-------------------------|
| 1. Enter system view.              | <b>system-view</b>                                                                                | N/A                     |
| 2. Configure DLDAP authentication. | <b>dldap authentication-mode</b> { <b>none</b>   { <b>md5</b>   <b>simple</b> } <i>password</i> } | <b>none</b> by default. |

## Resetting DLDAP state

After DLDAP detects a unidirectional link on a port, the port enters Disable state. In this case, DLDAP prompts you to shut down the port manually or it shuts down the port automatically depending on the user-defined port shutdown mode. To enable the port to perform DLDAP detect again, you can reset the DLDAP state of the port by using one of the following methods:

- If the port is shut down with the **shutdown** command manually, run the **undo shutdown** command on the port.
- If DLDAP automatically shuts down the port, run the **dldap reset** command on the port to enable the port to perform DLDAP detection again. Alternatively, you can wait for DLDAP to automatically enable the port when it detects that the link has been restored to bidirectional. For how to reset the DLDAP state by using the **dldap reset** command, see "[Resetting DLDAP state in system view](#)" and "[Resetting DLDAP state in interface view/port group view](#)."

The DLDAP state that the port transits to upon the DLDAP state reset operation depends on its physical state. If the port is physically down, it transits to Inactive state; if the port is physically up, it transits to Active state.

### Resetting DLDAP state in system view

Resetting DLDAP state in system view applies to all ports of the device.

To reset DLDAP in system view:

| Step                  | Command            |
|-----------------------|--------------------|
| 1. Enter system view. | <b>system-view</b> |
| 2. Reset DLDAP state. | <b>dldap reset</b> |

### Resetting DLDAP state in interface view/port group view

Resetting DLDAP state in interface view or port group view applies to the current port or all ports in the port group.

To reset DLDAP state in interface view/port group view:

| Step                                                         | Command                                                                                                                                                                        | Remarks                                                                                                                                                                                         |
|--------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Enter system view.                                        | <b>system-view</b>                                                                                                                                                             | N/A                                                                                                                                                                                             |
| 2. Enter Layer 2 Ethernet interface view or port group view. | Enter Layer 2 Ethernet interface view:<br><b>interface</b> <i>interface-type interface-number</i><br>Enter port group view:<br><b>port-group manual</b> <i>port-group-name</i> | Use either approach.<br>Configurations made in Layer 2 Ethernet interface view apply to the current port only. Configurations made in port group view apply to all the ports in the port group. |
| 3. Reset DLDAP state.                                        | <b>dldap reset</b>                                                                                                                                                             | N/A                                                                                                                                                                                             |

## Displaying and maintaining DLDAP

| Task                                                            | Command                                                                                                                                                       | Remarks                |
|-----------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|
| Display the DLDAP configuration of a port.                      | <b>display dldap</b> [ <i>interface-type interface-number</i> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]            | Available in any view  |
| Display the statistics on DLDAP packets passing through a port. | <b>display dldap statistics</b> [ <i>interface-type interface-number</i> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ] | Available in any view  |
| Clear the statistics on DLDAP packets passing through a port.   | <b>reset dldap statistics</b> [ <i>interface-type interface-number</i> ]                                                                                      | Available in user view |

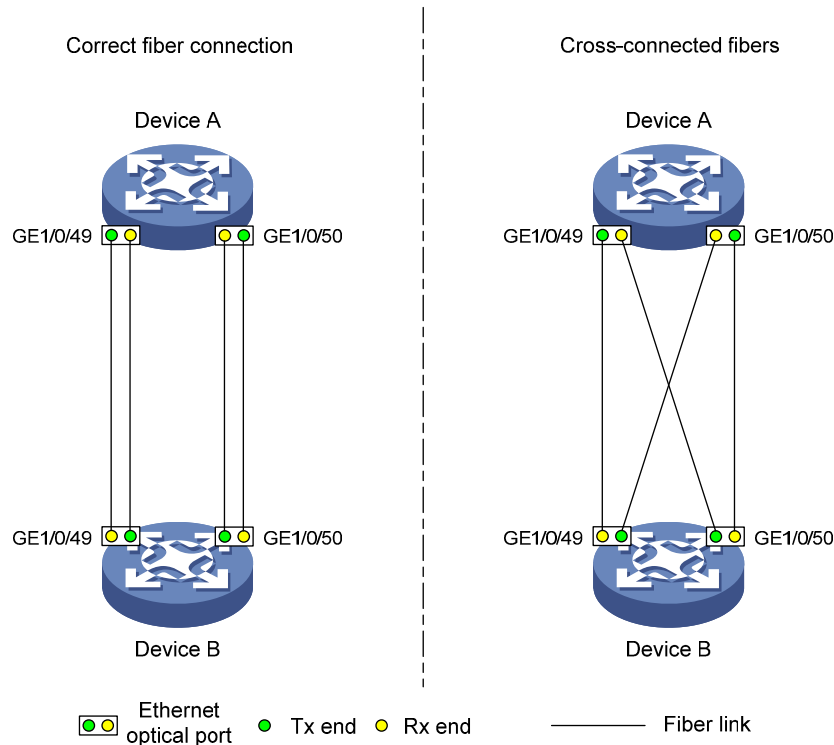
## DLDAP configuration examples

### Automatically shutting down unidirectional links

#### Network requirements

- As shown in [Figure 11](#), Device A and Device B are connected with two fiber pairs.
- Configure DLDAP to automatically shut down the faulty port upon detecting a unidirectional link, and automatically bring up the port after you clear the fault.

Figure 11 Network diagram



## Configuration procedure

### 1. Configure Device A:

# Enable DLDP globally.

```
<DeviceA> system-view
[DeviceA] dldp enable
```

# Configure GigabitEthernet 1/0/49 to operate in full duplex mode and at 1000 Mbps, and enable DLDP on the port.

```
[DeviceA] interface gigabitethernet 1/0/49
[DeviceA-GigabitEthernet1/0/49] duplex full
[DeviceA-GigabitEthernet1/0/49] speed 1000
[DeviceA-GigabitEthernet1/0/49] dldp enable
[DeviceA-GigabitEthernet1/0/49] quit
```

# Configure GigabitEthernet 1/0/50 to operate in full duplex mode and at 1000 Mbps, and enable DLDP on the port.

```
[DeviceA] interface gigabitethernet 1/0/50
[DeviceA-GigabitEthernet1/0/50] duplex full
[DeviceA-GigabitEthernet1/0/50] speed 1000
[DeviceA-GigabitEthernet1/0/50] dldp enable
[DeviceA-GigabitEthernet1/0/50] quit
```

# Set the DLDP mode to enhanced.

```
[DeviceA] dldp work-mode enhance
```

# Set the port shutdown mode to auto.

```
[DeviceA] dldp unidirectional-shutdown auto
```

### 2. Configure Device B:

# Enable DLDAP globally.

```
<DeviceB> system-view
[DeviceB] dldp enable
```

# Configure GigabitEthernet 1/0/49 to operate in full duplex mode and at 1000 Mbps, and enable DLDAP on it.

```
[DeviceB] interface gigabitethernet 1/0/49
[DeviceB-GigabitEthernet1/0/49] duplex full
[DeviceB-GigabitEthernet1/0/49] speed 1000
[DeviceB-GigabitEthernet1/0/49] dldp enable
[DeviceB-GigabitEthernet1/0/49] quit
```

# Configure GigabitEthernet 1/0/50 to operate in full duplex mode and at 1000 Mbps, and enable DLDAP on it.

```
[DeviceB] interface gigabitethernet 1/0/50
[DeviceB-GigabitEthernet1/0/50] duplex full
[DeviceB-GigabitEthernet1/0/50] speed 1000
[DeviceB-GigabitEthernet1/0/50] dldp enable
[DeviceB-GigabitEthernet1/0/50] quit
```

# Set the DLDAP mode to enhanced.

```
[DeviceB] dldp work-mode enhance
```

# Set the port shutdown mode to auto.

```
[DeviceB] dldp unidirectional-shutdown auto
```

### 3. Verify the configuration:

After the configurations are complete, you can use the **display dldp** command to display the DLDAP configuration information on ports.

# Display the DLDAP configuration information on all the DLDAP-enabled ports of Device A.

```
[DeviceA] display dldp
DLDP global status : enable
DLDP interval : 5s
DLDP work-mode : enhance
DLDP authentication-mode : none
DLDP unidirectional-shutdown : auto
DLDP delaydown-timer : 1s
The number of enabled ports is 2.
```

```
Interface GigabitEthernet1/0/49
DLDP port state : advertisement
DLDP link state : up
The neighbor number of the port is 1.
 Neighbor mac address : 0023-8956-3600
 Neighbor port index : 59
 Neighbor state : two way
 Neighbor aged time : 11
```

```
Interface GigabitEthernet1/0/50
DLDP port state : advertisement
DLDP link state : up
The neighbor number of the port is 1.
```

```
Neighbor mac address : 0023-8956-3600
Neighbor port index : 60
Neighbor state : two way
Neighbor aged time : 12
```

The output shows that both GigabitEthernet 1/0/49 and GigabitEthernet 1/0/50 are in Advertisement state, which means both links are bidirectional.

# Enable system information monitoring on Device A, and enable the display of log and trap information.

```
[DeviceA] quit
<DeviceA> terminal monitor
<DeviceA> terminal logging
<DeviceA> terminal trapping
```

The following log and trap information is displayed on Device A:

```
<DeviceA>
#Jan 18 17:36:18:798 2010 DeviceA DLDP/1/TrapOfUnidirectional: -Slot=1; Trap
1.3.6.1.4.1.25506.2.43.2.1.1 : DLDP detects a unidirectional link in port 17825792.

%Jan 18 17:36:18:799 2010 DeviceA IFNET/3/LINK_UPDOWN: GigabitEthernet1/0/49 link
status is DOWN.

%Jan 18 17:36:18:799 2010 DeviceA DLDP/3/DLDP_UNIDIRECTION_AUTO: -Slot=1; DLDP
detects a unidirectional link on port GigabitEthernet1/0/49. The transceiver has
malfunction in the Tx direction or cross-connected links exist between the local device
and its neighbor. The shutdown mode is AUTO. DLDP shuts down the port.

#Jan 18 17:36:20:189 2010 DeviceA DLDP/1/TrapOfUnidirectional: -Slot=1; Trap
1.3.6.1.4.1.25506.2.43.2.1.1 : DLDP detects a unidirectional link in port 17825793.

%Jan 18 17:36:20:189 2010 DeviceA IFNET/3/LINK_UPDOWN: GigabitEthernet1/0/50 link
status is DOWN.

%Jan 18 17:36:20:190 2010 DeviceA DLDP/3/DLDP_UNIDIRECTION_AUTO: -Slot=1; DLDP
detects a unidirectional link on port GigabitEthernet1/0/50. The transceiver has
malfunction in the Tx direction or cross-connected links exist between the local device
and its neighbor. The shutdown mode is AUTO. DLDP shuts down the port.

%Jan 15 16:54:56:040 2010 DeviceA DLDP/3/DLDP_UNIDIRECTION_AUTO_ENHANCE: -Slot=1; In
enhanced DLDP mode, port GigabitEthernet1/0/49 cannot detect its aged-out neighbor.
The transceiver has malfunction in the Tx direction or cross-connected links exist
between the local device and its neighbor. The shutdown mode is AUTO. DLDP shuts down
the port.
```

The output shows that the link status of both GigabitEthernet 1/0/49 and GigabitEthernet 1/0/50 is down, and DLDP has detected a unidirectional link on both ports and has automatically shut them down.

Assume that in this example, the unidirectional links are caused by cross-connected fibers. Correct the fiber connections on detecting the unidirectional link problem. As a result, the ports shut down by DLDP automatically recover, and Device A displays the following log information:

```
<DeviceA>
#Jan 18 17:47:33:869 2010 DeviceA IFNET/3/LINK_UPDOWN: GigabitEthernet1/0/49 link
status is UP.

#Jan 18 17:47:35:894 2010 DeviceA IFNET/3/LINK_UPDOWN: GigabitEthernet1/0/50 link
status is UP.
```



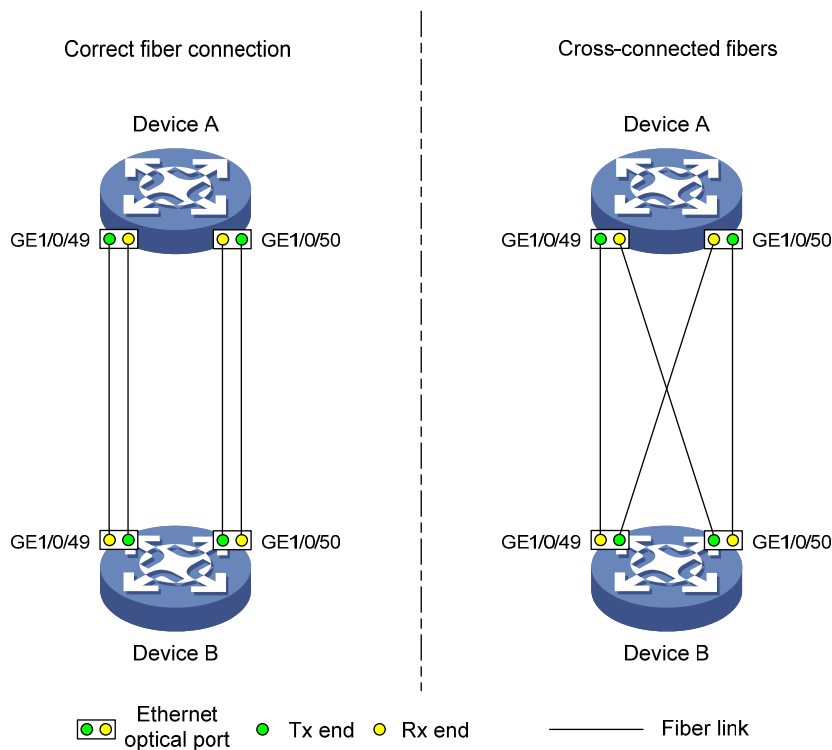
The output shows that the link status of both GigabitEthernet 1/0/49 and GigabitEthernet 1/0/50 is now up.

## Manually shutting down unidirectional links

### Network requirements

- As shown in [Figure 12](#), Device A and Device B are connected with two fiber pairs.
- Configure DLDP to send information when a unidirectional link is detected, to remind the network administrator to manually shut down the faulty port.

**Figure 12 Network diagram**



### Configuration procedure

#### 1. Configure Device A:

# Enable DLDP globally.

```
<DeviceA> system-view
[DeviceA] dldp enable
```

# Configure GigabitEthernet 1/0/49 to operate in full duplex mode and at 1000 Mbps, and enable DLDP on the port.

```
[DeviceA] interface gigabitethernet 1/0/49
[DeviceA-GigabitEthernet1/0/49] duplex full
[DeviceA-GigabitEthernet1/0/49] speed 1000
[DeviceA-GigabitEthernet1/0/49] dldp enable
[DeviceA-GigabitEthernet1/0/49] quit
```

# Configure GigabitEthernet 1/0/50 to operate in full duplex mode and at 1000 Mbps, and enable DLDP on the port.

```
[DeviceA] interface gigabitethernet 1/0/50
```

```
[DeviceA-GigabitEthernet1/0/50] duplex full
[DeviceA-GigabitEthernet1/0/50] speed 1000
[DeviceA-GigabitEthernet1/0/50] dldp enable
[DeviceA-GigabitEthernet1/0/50] quit
```

# Set the DLDP mode to enhanced.

```
[DeviceA] dldp work-mode enhance
```

# Set the port shutdown mode to manual.

```
[DeviceA] dldp unidirectional-shutdown manual
```

## 2. Configure Device B:

# Enable DLDP globally.

```
<DeviceB> system-view
```

```
[DeviceB] dldp enable
```

# Configure GigabitEthernet 1/0/49 to operate in full duplex mode and at 1000 Mbps, and enable DLDP on it.

```
[DeviceB] interface gigabitethernet 1/0/49
[DeviceB-GigabitEthernet1/0/49] duplex full
[DeviceB-GigabitEthernet1/0/49] speed 1000
[DeviceB-GigabitEthernet1/0/49] dldp enable
[DeviceB-GigabitEthernet1/0/49] quit
```

# Configure GigabitEthernet 1/0/50 to operate in full duplex mode and at 1000 Mbps, and enable DLDP on it.

```
[DeviceB] interface gigabitethernet 1/0/50
[DeviceB-GigabitEthernet1/0/50] duplex full
[DeviceB-GigabitEthernet1/0/50] speed 1000
[DeviceB-GigabitEthernet1/0/50] dldp enable
[DeviceB-GigabitEthernet1/0/50] quit
```

# Set the DLDP mode to enhanced.

```
[DeviceB] dldp work-mode enhance
```

# Set the port shutdown mode to manual.

```
[DeviceB] dldp unidirectional-shutdown manual
```

## 3. Verify the configuration:

After the configurations are complete, you can use the **display dldp** command to display the DLDP configuration information on ports.

# Display the DLDP configuration information on all the DLDP-enabled ports of Device A.

```
[DeviceA] display dldp
DLDP global status : enable
DLDP interval : 5s
DLDP work-mode : enhance
DLDP authentication-mode : none
DLDP unidirectional-shutdown : manual
DLDP delaydown-timer : 1s
The number of enabled ports is 2.
```

```
Interface GigabitEthernet1/0/49
DLDP port state : advertisement
DLDP link state : up
```

```
The neighbor number of the port is 1.
 Neighbor mac address : 0023-8956-3600
 Neighbor port index : 59
 Neighbor state : two way
 Neighbor aged time : 11
```

```
Interface GigabitEthernet1/0/50
 DLDP port state : advertisement
 DLDP link state : up
 The neighbor number of the port is 1.
 Neighbor mac address : 0023-8956-3600
 Neighbor port index : 60
 Neighbor state : two way
 Neighbor aged time : 12
```

The output shows that both GigabitEthernet 1/0/49 and GigabitEthernet 1/0/50 are in Advertisement state, which means both links are bidirectional.

# Enable system information monitoring on Device A, and enable the display of log and trap information.

```
[DeviceA] quit
<DeviceA> terminal monitor
<DeviceA> terminal logging
<DeviceA> terminal trapping
```

The following log and trap information is displayed on Device A:

```
<DeviceA>
#Jan 18 18:10:38:481 2010 DeviceA DLDP/1/TrapOfUnidirectional: -Slot=1; Trap
1.3.6.1.4.1.25506.2.43.2.1.1 : DLDP detects a unidirectional link in port 17825792.
```

```
%Jan 18 18:10:38:481 2010 DeviceA DLDP/3/DLDP_UNIDIRECTION_MANUAL: -Slot=1; DLDP
detects a unidirectional link on port GigabitEthernet1/0/49. The transceiver has
malfunction in the Tx direction or cross-connected links exist between the local device
and its neighbor. The shutdown mode is MANUAL. The port needs to be shut down by the
user.
```

```
#Jan 18 18:10:38:618 2010 DeviceA DLDP/1/TrapOfUnidirectional: -Slot=1; Trap
1.3.6.1.4.1.25506.2.43.2.1.1 : DLDP detects a unidirectional link in port 17825793.
```

```
%Jan 18 18:10:38:618 2010 DeviceA DLDP/3/DLDP_UNIDIRECTION_MANUAL: -Slot=1; DLDP
detects a unidirectional link on port GigabitEthernet1/0/50. The transceiver has
malfunction in the Tx direction or cross-connected links exist between the local device
and its neighbor. The shutdown mode is MANUAL. The port needs to be shut down by the
user.
```

The output shows that DLDP has detected a unidirectional link on both GigabitEthernet 1/0/49 and GigabitEthernet 1/0/50, and is asking you to shut down the faulty ports manually.

After you shut down GigabitEthernet 1/0/49 and GigabitEthernet 1/0/50, the following log information is displayed:

```
<DeviceA> system-view
[DeviceA] interface gigabitethernet 1/0/49
[DeviceA-GigabitEthernet1/0/49] shutdown
%Jan 18 18:16:12:044 2010 DeviceA IFNET/3/LINK_UPDOWN: GigabitEthernet1/0/49 link
status is DOWN.
```

```
[DeviceA-GigabitEthernet1/0/49] quit
[DeviceA] interface gigabitethernet 1/0/50
[DeviceA-GigabitEthernet1/0/50] shutdown
%Jan 18 18:18:03:583 2010 DeviceA IFNET/3/LINK_UPDOWN: GigabitEthernet1/0/50 link
status is DOWN.
```

The output shows that the link status of both GigabitEthernet 1/0/49 and GigabitEthernet 1/0/50 is down.

Assume that in this example, the unidirectional links are caused by cross-connected fibers. Correct the fiber connections, and then bring up the ports shut down earlier.

# On Device A, bring up GigabitEthernet 1/0/49 and GigabitEthernet 1/0/50:

```
[DeviceA-GigabitEthernet1/0/50] undo shutdown
[DeviceA-GigabitEthernet1/0/50]
%Jan 18 18:22:11:698 2010 DeviceA IFNET/3/LINK_UPDOWN: GigabitEthernet1/0/50 link
status is UP.
[DeviceA-GigabitEthernet1/0/50] quit
[DeviceA] interface gigabitethernet 1/0/49
[DeviceA-GigabitEthernet1/0/49] undo shutdown
[DeviceA-GigabitEthernet1/0/49]
%Jan 18 18:22:46:065 2010 DeviceA IFNET/3/LINK_UPDOWN: GigabitEthernet1/0/49 link
status is UP.
```

The output shows that the link status of both GigabitEthernet 1/0/49 and GigabitEthernet 1/0/50 is now up.

## Troubleshooting DLDAP

### Symptom

Two DLDAP-enabled devices, Device A and Device B, are connected through two fiber pairs, in which two fibers are cross-connected. The unidirectional links cannot be detected; all the four ports involved are in Advertisement state.

### Analysis

The problem can be caused by the following.

- The intervals to send Advertisement packets on Device A and Device B are not the same.
- DLDAP authentication modes/passwords on Device A and Device B are not the same.

### Solution

Make sure the interval to send Advertisement packets, the authentication mode, and the password configured on Device A and Device B are the same.

# Configuring RRPP

## RRPP overview

The Rapid Ring Protection Protocol (RRPP) is a link layer protocol designed for Ethernet rings. RRPP can prevent broadcast storms caused by data loops when an Ethernet ring is healthy, and rapidly restore the communication paths between the nodes in the event that a link is disconnected on the ring.

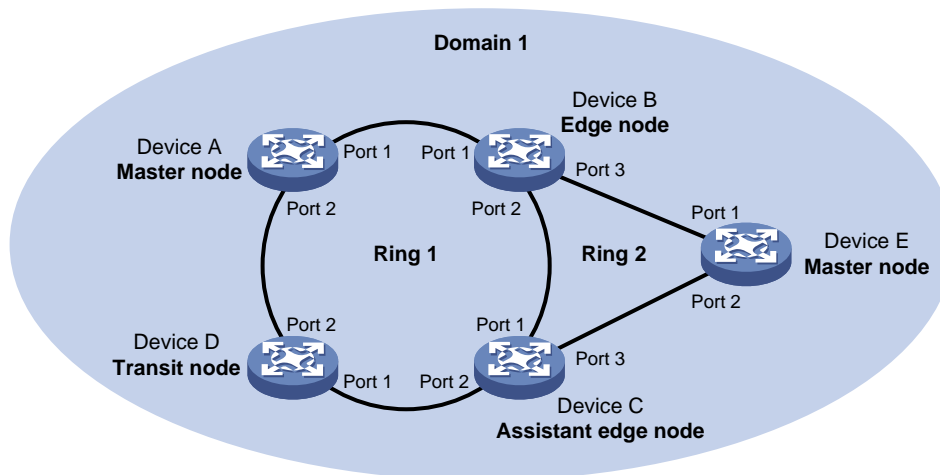
## Background

Metropolitan area networks (MANs) and enterprise networks usually use the ring structure to improve reliability. However, services will be interrupted if any node in the ring network fails. A ring network usually uses Resilient Packet Ring (RPR) or Ethernet rings. RPR is high in cost because it needs dedicated hardware. Contrarily, the Ethernet ring technology is more mature and economical, so it is increasingly widely used in MANs and enterprise networks.

Rapid Spanning Tree Protocol (RSTP), Per VLAN Spanning Tree (PVST), Multiple Spanning Tree Protocol (MSTP), and RRPP can eliminate Layer-2 loops. RSTP, PVST, and MSTP are mature. However, they take several seconds to converge. RRPP is an Ethernet ring-specific data link layer protocol, and it converges faster than RSTP, PVST, and MSTP. Additionally, the convergence time of RRPP is independent of the number of nodes in the Ethernet ring. RRPP can be applied to large-diameter networks.

## Basic concepts in RRPP

Figure 13 RRPP networking diagram



### RRPP domain

The interconnected devices with the same domain ID and control VLANs constitute an RRPP domain. An RRPP domain contains the following elements—primary ring, subring, control VLAN, master node, transit node, primary port, secondary port, common port, edge port, and so on.

As shown in Figure 13, Domain 1 is an RRPP domain, including two RRPP rings: Ring 1 and Ring 2. All the nodes on the two RRPP rings belong to the RRPP domain.

## RRPP ring

A ring-shaped Ethernet topology is called an "RRPP ring". RRPP rings fall into two types: primary ring and subring. You can configure a ring as either the primary ring or a subring by specifying its ring level. The primary ring is of level 0, and a subring is of level 1. An RRPP domain contains one or multiple RRPP rings, one serving as the primary ring and the others serving as subrings. A ring can be in one of the following states:

- **Health state**—All the physical links on the Ethernet ring are connected
- **Disconnect state**—Some physical links on the Ethernet ring are broken

As shown in [Figure 13](#), Domain 1 contains two RRPP rings: Ring 1 and Ring 2. The level of Ring 1 is set to 0, and that of Ring 2 is set to 1. Ring 1 is configured as the primary ring, and Ring 2 is configured as a subring.

## Control VLAN and data VLAN

### 1. Control VLAN

In an RRPP domain, a control VLAN is a VLAN dedicated to transferring Rapid Ring Protection Protocol Data Units (RRPPDUs). On a device, the ports accessing an RRPP ring belong to the control VLANs of the ring, and only such ports can join the control VLANs.

An RRPP domain is configured with two control VLANs: one primary control VLAN, which is the control VLAN for the primary ring, and one secondary control VLAN, which is the control VLAN for subrings. All subrings in the same RRPP domain share the same secondary control VLAN. After you specify a VLAN as the primary control VLAN, the system automatically configures the VLAN whose ID is the primary control VLAN ID plus one as the secondary control VLAN.

IP address configuration is prohibited on the control VLAN interfaces.

### 2. Data VLAN

A data VLAN is a VLAN dedicated to transferring data packets. Both RRPP ports and non-RRPP ports can be assigned to a data VLAN.

## Node

Each device on an RRPP ring is a node. The role of a node is configurable. RRPP has the following node roles:

- **Master node**—Each ring has one and only one master node. The master node initiates the polling mechanism and determines the operations to be performed after a change in topology.
- **Transit node**—Transit nodes include all the nodes except the master node on the primary ring and all the nodes on subrings except the master nodes and the nodes where the primary ring intersects with the subrings. A transit node monitors the state of its directly-connected RRPP links and notifies the master node of the link state changes, if any. Based on the link state changes, the master node decides the operations to be performed.
- **Edge node**—A node residing on both the primary ring and a subring at the same time. An edge node is a special transit node that serves as a transit node on the primary ring and an edge node on the subring.
- **Assistant-edge node**—A node residing on both the primary ring and a subring at the same time. An assistant-edge node is a special transit node that serves as a transit node on the primary ring and an assistant-edge node on the subring. This node works in conjunction with the edge node to detect the integrity of the primary ring and to perform loop guard.

As shown in [Figure 13](#), Ring 1 is the primary ring and Ring 2 is a subring. Device A is the master node of Ring 1, and Device B, Device C, and Device D are the transit nodes of Ring 1. Device E is the master node of Ring 2, Device B is the edge node of Ring 2, and Device C is the assistant-edge node of Ring 2.

## Primary port and secondary port

Each master node or transit node has two ports connected to an RRPP ring, one serving as the primary port and the other serving as the secondary port. You can determine the port's role.

1. In terms of functionality, the primary port and the secondary port of a master node have the following differences:
  - The primary port and the secondary port are designed to play the role of sending and receiving loop-detect packets respectively.
  - When an RRPP ring is in Health state, the secondary port of the master node will logically deny data VLANs and permit only the packets of the control VLANs.
  - When an RRPP ring is in Disconnect state, the secondary port of the master node will permit data VLANs (forward packets of data VLANs).
2. In terms of functionality, the primary port and the secondary port of a transit node have no difference. Both are designed for transferring protocol packets and data packets over an RRPP ring.

As shown in [Figure 13](#), Device A is the master node of Ring 1. Port 1 and Port 2 are the primary port and the secondary port of the master node on Ring 1 respectively. Device B, Device C, and Device D are the transit nodes of Ring 1. Their Port 1 and Port 2 are the primary port and the secondary port on Ring 1 respectively.

## Common port and edge port

The ports connecting the edge node and assistant-edge node to the primary ring are common ports. The ports connecting the edge node and assistant-edge node only to the subrings are edge ports.

As shown in [Figure 13](#), Device B and Device C lie on Ring 1 and Ring 2. Device B's Port 1 and Port 2 and Device C's Port 1 and Port 2 access the primary ring, so they are common ports. Device B's Port 3 and Device C's Port 3 access only the subring, so they are edge ports.

## RRPP ring group

To reduce Edge-Hello traffic, you can configure a group of subrings on the edge node or assistant-edge node. For more information about Edge-Hello packets, see "[RRPPDU](#)." You must configure a device as the edge node of these subrings, and another device as the assistant-edge node of these subrings. Additionally, the subrings of the edge node and assistant-edge node must connect to the same subring packet tunnels in major ring (SRPTs) so that Edge-Hello packets of the edge node of these subrings travel to the assistant-edge node of these subrings over the same link.

An RRPP ring group configured on the edge node is an edge node RRPP ring group, and an RRPP ring group configured on an assistant-edge node is an assistant-edge node RRPP ring group. Up to one subring in an edge node RRPP ring group is allowed to send Edge-Hello packets.

# RRPPDU

**Table 18 RRPPDU types and their functions**

Type	Description
Hello	The master node initiates Hello packets to detect the integrity of a ring in a network.
Link-Down	The transit node, the edge node, or the assistant-edge node initiates Link-Down packets to notify the master node of the disappearance of a ring in case of a link failure.

Type	Description
Common-Flush-FDB	The master node initiates Common-Flush-FDB (FDB stands for Forwarding Database) packets to instruct the transit nodes to update their own MAC entries and ARP/ND entries when an RRPP ring transits to Disconnect state.
Complete-Flush-FDB	The master node initiates Complete-Flush-FDB packets to instruct the transit nodes to update their own MAC entries and ARP/ND entries and release blocked ports from being blocked temporarily when an RRPP ring transits to Health state.
Edge-Hello	The edge node initiates Edge-Hello packets to examine the SRPTs between the edge node and the assistant-edge node.
Major-Fault	The assistant-edge node initiates Major-Fault packets to notify the edge node of SRPT failure when an SRPT between edge node and assistant-edge node is torn down.

**NOTE:**

RRPPDUs of subrings are transmitted as data packets in the primary ring, and RRPPDUs of the primary ring can only be transmitted within the primary ring.

## RRPP timers

When RRPP checks the link state of an Ethernet ring, the master node sends Hello packets out of the primary port according to the Hello timer and determines whether its secondary port receives the Hello packets based on the Fail timer.

- The Hello timer specifies the interval at which the master node sends Hello packets out of the primary port.
- The Fail timer specifies the maximum delay between the master node sending Hello packets out of the primary port and the secondary port receiving the Hello packets from the primary port. If the secondary port receives the Hello packets sent by the local master node before the Fail timer expires, the overall ring is in Health state. Otherwise, the ring transits into the Disconnect state.

**NOTE:**

In an RRPP domain, a transit node learns the Hello timer value and the Fail timer value on the master node through the received Hello packets, ensuring that all nodes in the ring network are consistent in the two timer settings.

## How RRPP works

### Polling mechanism

The polling mechanism is used by the master node of an RRPP ring to check the Health state of the ring network.

The master node periodically sends Hello packets out of its primary port, and these Hello packets travel through each transit node on the ring in turn:

- If the ring is complete, the secondary port of the master node will receive Hello packets before the Fail timer expires and the master node will keep the secondary port blocked.
- If the ring is torn down, the secondary port of the master node will fail to receive Hello packets before the Fail timer expires. The master node will release the secondary port from blocking data



VLANs and sending Common-Flush-FDB packets to instruct all transit nodes to update their own MAC entries and ARP/ND entries.

### Link down alarm mechanism

The transit node, the edge node or the assistant-edge node sends Link-Down packets to the master node immediately when they find any of its own ports belonging to an RRPP domain are down. Upon the receipt of a Link-Down packet, the master node releases the secondary port from blocking data VLANs and sending Common-Flush-FDB packet to instruct all the transit nodes, the edge nodes, and the assistant-edge nodes to update their own MAC entries and ARP/ND entries. After each node updates its own entries, traffic is switched to the normal link.

### Ring recovery

The master node may find that the ring is restored after a period of time after the ports belonging to the RRPP domain on the transit nodes, the edge nodes, or the assistant-edge nodes are brought up again. A temporary loop may arise in the data VLAN during this period. As a result, broadcast storm occurs.

To prevent temporary loops, non-master nodes block them immediately (and permit only the packets of the control VLAN to pass through) when they find their ports accessing the ring are brought up again. The blocked ports are activated only when the nodes are sure that no loop will be brought forth by these ports.

### Broadcast storm suppression mechanism in a multi-homed subring in case of SRPT failure

As shown in [Figure 17](#), Ring 1 is the primary ring, and Ring 2 and Ring 3 are subrings. When the two SRPTs between the edge node and the assistant-edge node are down, the master nodes of Ring 2 and Ring 3 will open their respective secondary ports, generating a loop among Device B, Device C, Device E, and Device F. As a result, a broadcast storm occurs.

To prevent generating this loop, the edge node will block the edge port temporarily. The blocked edge port is activated only when the edge node is sure that no loop will be brought forth when the edge port is activated.

### Load balancing

In a ring network, maybe traffic of multiple VLANs is transmitted at the same time. RRPP can implement load balancing for the traffic by transmitting traffic of different VLANs along different paths.

By configuring an individual RRPP domain for transmitting the traffic of the specified VLANs (protected VLANs) in a ring network, traffic of different VLANs can be transmitted according to different topologies in the ring network. In this way, load balancing is achieved.

As shown in [Figure 18](#), Ring 1 is configured as the primary ring of Domain 1 and Domain 2, which are configured with different protected VLANs. Device A is the master node of Ring 1 in Domain 1, and Device B is the master node of Ring 1 in Domain 2. With such configurations, traffic of different VLANs can be transmitted on different links to achieve load balancing in the single-ring network.

### RRPP ring group

In an edge node RRPP ring group, only an activated subring with the lowest domain ID and ring ID can send Edge-Hello packets. In an assistant-edge node RRPP ring group, any activated subring that has received Edge-Hello packets will forward these packets to the other activated subrings. With an edge node RRPP ring group and an assistant-edge node RRPP ring group configured, only one subring sends Edge-Hello packets on the edge node, and only one subring receives Edge-Hello packets on the assistant-edge node, reducing CPU workload.

As shown in [Figure 17](#), Device B is the edge node of Ring 2 and Ring 3, and Device C is the assistant-edge node of Ring 2 and Ring 3. Device B and Device C must send or receive Edge-Hello

packets frequently. If more subrings are configured or if load balancing is configured for multiple domains, Device B and Device C will send or receive a mass of Edge-Hello packets.

To reduce Edge-Hello traffic, you can assign Ring 2 and Ring 3 to an RRPP ring group configured on the edge node Device B and assign Ring 2 and Ring 3 to an RRPP ring group configured on Device C. After such configurations, if all rings are activated, only Ring 2 on Device B sends Edge-Hello packets.

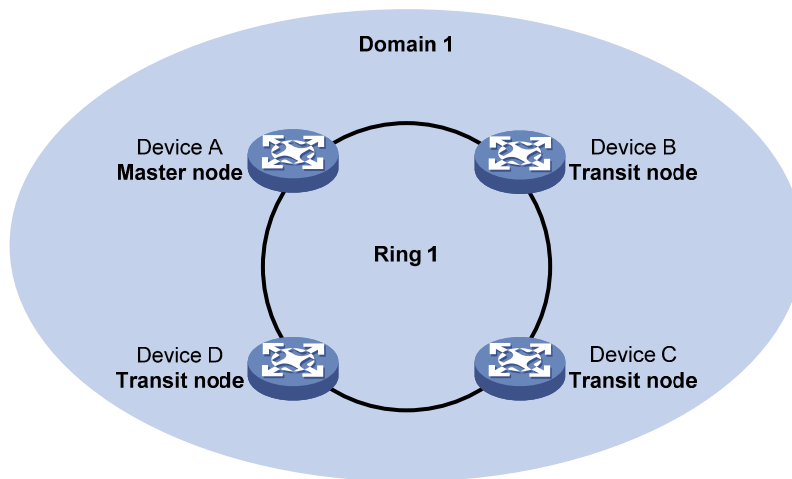
## Typical RRPP networking

Here are several typical networking applications.

### Single ring

As shown in [Figure 14](#), only a single ring exists in the network topology. You only need to define an RRPP domain.

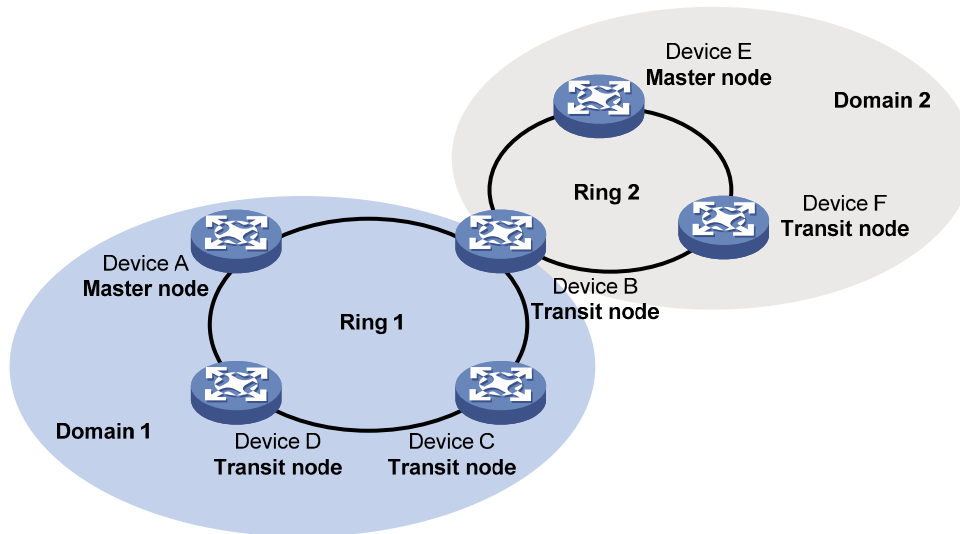
**Figure 14 Schematic diagram for a single-ring network**



### Tangent rings

As shown in [Figure 15](#), two or more rings are in the network topology and only one common node exists between rings. You must define an RRPP domain for each ring.

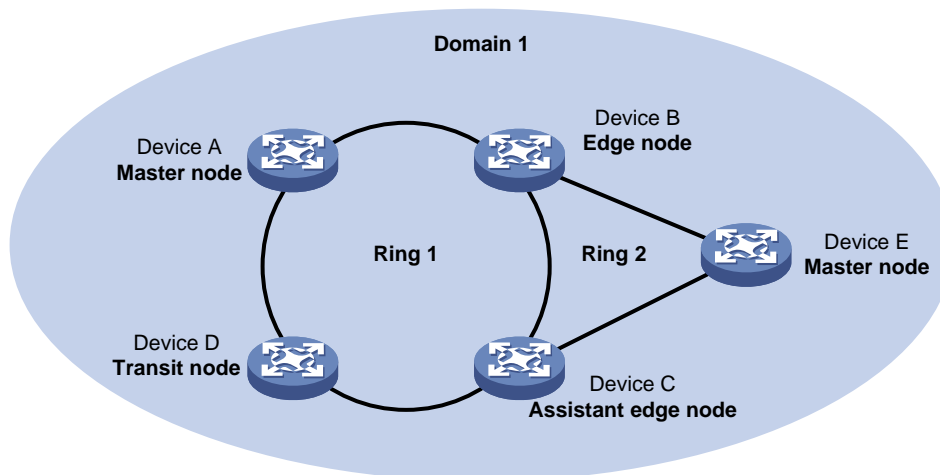
Figure 15 Schematic diagram for a tangent-ring network



### Intersecting rings

As shown in Figure 16, two or more rings are in the network topology and two common nodes exist between rings. You only need to define an RRPP domain and configure one ring as the primary ring and the other rings as subrings.

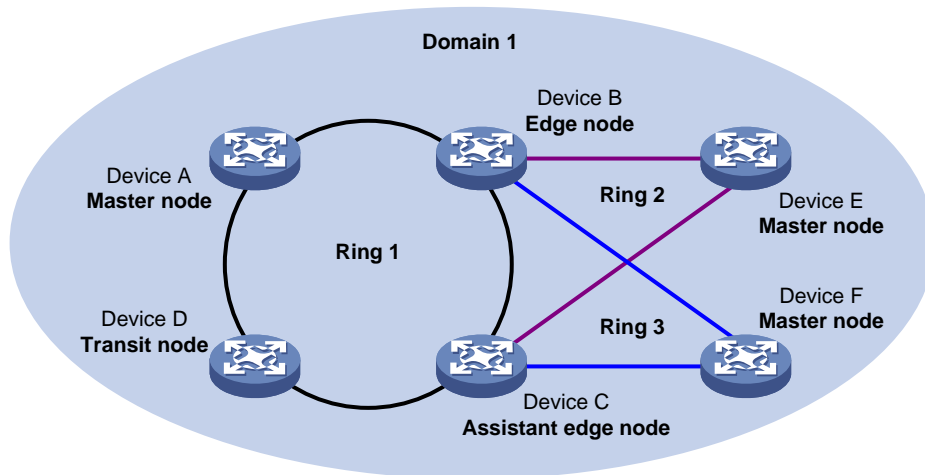
Figure 16 Schematic diagram for an intersecting-ring network



### Dual homed rings

As shown in Figure 17, two or more rings are in the network topology and two similar common nodes exist between rings. You only need to define an RRPP domain and configure one ring as the primary ring and the other rings as subrings.

**Figure 17 Schematic diagram for a dual-homed-ring network**

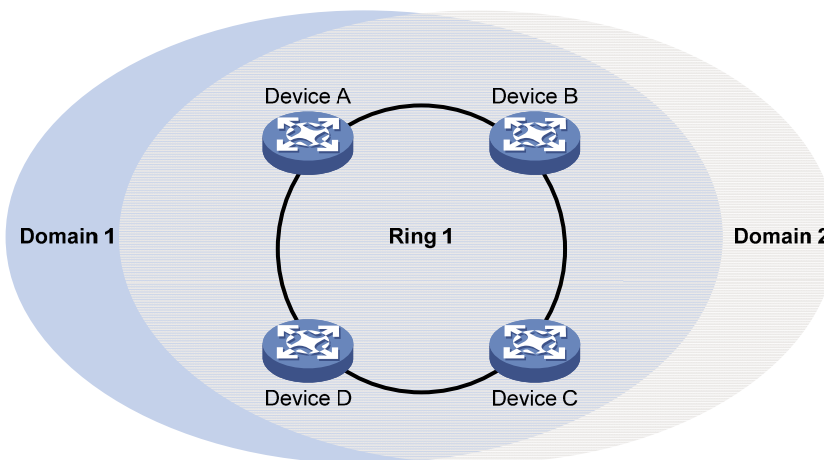


### Single-ring load balancing

In a single-ring network, you can achieve load balancing by configuring multiple domains.

As shown in Figure 18, Ring 1 is configured as the primary ring of both Domain 1 and Domain 2. Domain 1 and Domain 2 are configured with different protected VLANs. In Domain 1, Device A is configured as the master node of Ring 1. In Domain 2, Device B is configured as the master node of Ring 1. Such configurations enable the ring to block different links based on VLANs, and single-ring load balancing is achieved.

**Figure 18 Schematic diagram for a single-ring load balancing network**

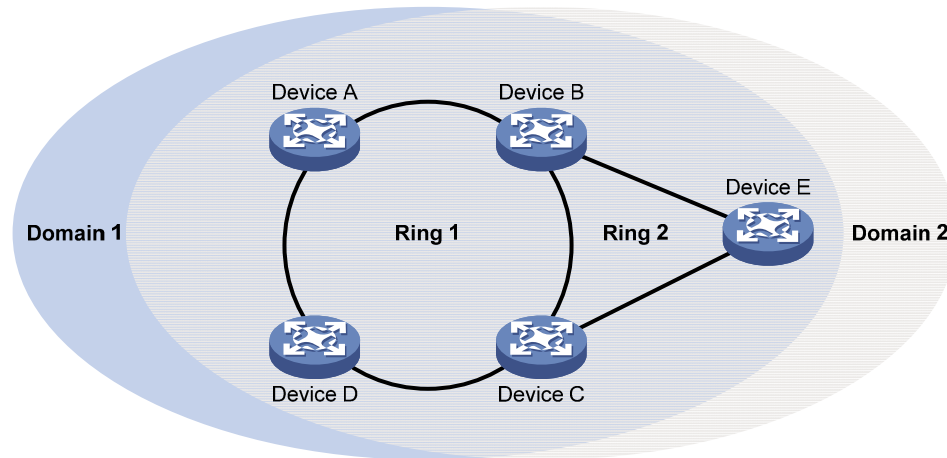


### Intersecting-ring load balancing

In an intersecting-ring network, you can also achieve load balancing by configuring multiple domains.

As shown in Figure 19, Ring 1 is the primary ring, and Ring 2 is the subring in both Domain 1 and Domain 2. Domain 1 and Domain 2 are configured with different protected VLANs. Device A is configured as the master node of Ring 1 in Domain 1. Device D is configured as the master node of Ring 1 in Domain 2. Device E is configured as the master node of Ring 2 in both Domain 1 and Domain 2. However, different ports on Device E are blocked in Domain 1 and Domain 2. With the configurations, you can enable traffic of different VLANs to travel over different paths in the subring and primary ring to achieve intersecting-ring load balancing.

Figure 19 Schematic diagram for an intersecting-ring load balancing network



## Protocols and standards

RFC 3619 *Extreme Networks' Ethernet Automatic Protection Switching (EAPS) Version 1* is related to RRPP.

## RRPP configuration task list

You can create RRPP domains based on service planning, specify control VLANs and data VLANs for each RRPP domain, and then determine the ring roles and node roles based on the traffic paths in each RRPP domain.

Complete the following tasks to configure RRPP:

Task	Remarks
Creating an RRPP domain	Required Perform this task on all nodes in the RRPP domain.
Configuring control VLANs	Required Perform this task on all nodes in the RRPP domain.
Configuring protected VLANs	Required Perform this task on all nodes in the RRPP domain.
Configuring RRPP rings	Configuring RRPP ports Required Perform this task on all nodes in the RRPP domain.
	Configuring RRPP nodes Required Perform this task on all nodes in the RRPP domain.
Activating an RRPP domain	Required Perform this task on all nodes in the RRPP domain.
Configuring RRPP timers	Optional Perform this task on the master node in the RRPP domain.

Task	Remarks
Configuring an RRPP ring group	Optional Perform this task on the edge node and assistant-edge node in the RRPP domain.

**NOTE:**

- RRPP does not have an auto election mechanism, so you must configure each node in the ring network properly for RRPP to monitor and protect the ring network.
- Before configuring RRPP, you must construct a ring-shaped Ethernet topology physically.

## Creating an RRPP domain

When creating an RRPP domain, specify a domain ID, which uniquely identifies an RRPP domain. All devices in the same RRPP domain must be configured with the same domain ID.

Perform this configuration on devices you want to configure as nodes in the RRPP domain.

To create an RRPP domain:

Step	Command
1. Enter system view.	<b>system-view</b>
2. Create an RRPP domain, and enter RRPP domain view.	<b>rrpp domain <i>domain-id</i></b>

## Configuring control VLANs

Before configuring RRPP rings in an RRPP domain, configure the same control VLANs for all nodes in the RRPP domain first.

Perform this configuration on all nodes in the RRPP domain to be configured.

## Configuration guidelines

- When you configure existing VLANs as control VLANs, the system prompts errors.
- To ensure proper forwarding of RRPPDUs, do not enable 802.1Q in 802.1Q (QinQ) or VLAN mapping on the control VLANs.
- To make sure RRPPDUs can be sent and received correctly, do not configure the default VLAN of a port accessing an RRPP ring as the primary control VLAN or the secondary control VLAN.
- To transparently transmit RRPPDUs on a device not configured with RRPP, you must ensure only the two ports connecting the device to the RRPP ring permit the packets of the control VLANs. Otherwise, the packets from other VLANs may go into the control VLANs in transparent transmission mode and strike the RRPP ring.

## Configuration procedure

To configure control VLANs:

Step	Command
1. Enter system view.	<b>system-view</b>
2. Enter RRPP domain view.	<b>rrpp domain</b> <i>domain-id</i>
3. Configure the primary control VLAN for the RRPP domain.	<b>control-vlan</b> <i>vlan-id</i>

## Configuring protected VLANs

Before configuring RRPP rings in an RRPP domain, configure the same protected VLANs for all nodes in the RRPP domain first. All VLANs that the RRPP ports are assigned to should be protected by the RRPP domains.

You can configure protected VLANs through referencing Multiple Spanning Tree Instances (MSTIs). Before configuring protected VLANs, configure the mappings between MSTIs and the VLANs to be protected (a device working in PVST mode automatically maps VLANs to MSTIs). For more information about MSTIs and PVST, see *Layer 2—LAN Switching Configuration Guide*.

Perform this configuration on all nodes in the RRPP domain to be configured.

To configure protected VLANs:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Enter MST region view.	<b>stp region-configuration</b>	Not required if the device is operating in PVST mode. For more information about the command, see <i>Layer 2—LAN Switching Command Reference</i> .
3. Configure the VLAN-to-instance mapping table.	Approach 1: <b>instance</b> <i>instance-id</i> <b>vlan</b> <i>vlan-list</i> Approach 2: <b>vlan-mapping modulo</b> <i>modulo</i>	Optional. Use either approach. All VLANs in an MST region are mapped to MSTI 0 (the CIST) by default. Not required if the device is operating in PVST mode. For more information about the commands, see <i>Layer 2—LAN Switching Command Reference</i> .
4. Activate MST region configuration manually.	<b>active region-configuration</b>	Not required if the device is operating in PVST mode. For more information about the command, see <i>Layer 2—LAN Switching Command Reference</i> .

Step	Command	Remarks
5.	Display the currently activated configuration information of the MST region. <b>display stp region-configuration</b> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]	Optional. Available in any view. The command output includes VLAN-to-instance mappings. For more information about the command, see <i>Layer 2—LAN Switching Command Reference</i> .
6.	Return to system view. <b>quit</b>	Not required if the device is operating in PVST mode.
7.	Enter RRPP domain view. <b>rrpp domain</b> <i>domain-id</i>	N/A
8.	Configure protected VLANs for the RRPP domain. <b>protected-vlan reference-instance</b> <i>instance-id-list</i>	By default, no protected VLAN is configured for an RRPP domain.

**NOTE:**

When configuring load balancing, you must configure different protected VLANs for different RRPP domains.

## Configuring RRPP rings

When configuring an RRPP ring, you must make some configurations on the ports connecting each node to the RRPP ring before configuring the nodes.

RRPP ports (connecting devices to an RRPP ring) must be Layer-2 Ethernet ports or Layer-2 aggregate interfaces and cannot be member ports of any aggregation group, service loopback group, or smart link group.

After configuring a Layer-2 aggregate interface as an RRPP port, you can still assign ports to or remove ports from the aggregation group corresponding to the interface.

## Configuring RRPP ports

Perform this configuration on each node's ports intended for accessing RRPP rings.

### Configuration guidelines

- RRPP ports always allow packets of the control VLANs to pass through.
- For more information about the **port link-type trunk**, **port trunk permit vlan**, and **undo stp enable** commands, see *Layer 2—LAN Switching Command Reference*.
- The 802.1p priority of trusted packets on the RRPP ports must be configured, so that RRPP packets take higher precedence than data packets when passing through the RRPP ports. For more information about the **qos trust dot1p** command, see *ACL and QoS Command Reference*.
- Do not enable OAM remote loopback function on an RRPP port. Otherwise, it may cause a temporary broadcast storm.
- Do not configure a port accessing an RRPP ring as the destination port of a mirroring group.
- Do not configure physical-link-state change suppression time on a port accessing an RRPP ring to accelerate topology convergence. For more information, see the **undo link-delay** command (*Layer 2—LAN Switching Command Reference*).



## Configuration procedure

To configure RRPP ports:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Enter Layer 2 Ethernet interface view or Layer 2 aggregation interface view.	<b>interface</b> <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the link type of the interface as trunk.	<b>port link-type trunk</b>	By default, the link type of an interface is access.
4. Assign the trunk port to the protected VLANs of the RRPP domain.	<b>port trunk permit vlan</b> { <i>vlan-id-list</i>   <b>all</b> }	By default, a trunk port allows only packets of VLAN 1 to pass through.
5. Disable the spanning tree feature.	<b>undo stp enable</b>	Enabled by default.
6. Configure the port to trust the 802.1p precedence of the received packets.	<b>qos trust dot1p</b>	By default, the port priority is trusted.

## Configuring RRPP nodes

If a device carries multiple RRPP rings in an RRPP domain, only one ring can be configured as the primary ring on the device, and the role of the device on a subring can only be an edge node or an assistant-edge node.

### Specifying a master node

Perform this configuration on a device to be configured as a master node.

To specify a master node:

Step	Command
1. Enter system view.	<b>system-view</b>
2. Enter RRPP domain view.	<b>rrpp domain</b> <i>domain-id</i>
3. Specify the current device as the master node of the ring, and specify the primary port and the secondary port.	<b>ring</b> <i>ring-id</i> <b>node-mode master</b> [ <b>primary-port</b> <i>interface-type interface-number</i> ] [ <b>secondary-port</b> <i>interface-type interface-number</i> ] <b>level</b> <i>level-value</i>

### Specifying a transit node

Perform this configuration on a device to be configured as a transit node.

To specify a transit node:

Step	Command
1. Enter system view.	<b>system-view</b>
2. Enter RRPP domain view.	<b>rrpp domain</b> <i>domain-id</i>
3. Specify the current device as a transit node of the ring, and specify the primary port and the secondary port.	<b>ring</b> <i>ring-id</i> <b>node-mode transit</b> [ <b>primary-port</b> <i>interface-type interface-number</i> ] [ <b>secondary-port</b> <i>interface-type interface-number</i> ] <b>level</b> <i>level-value</i>

## Specifying an edge node

When configuring an edge node, you must first configure the primary ring before configuring the subrings.

Perform this configuration on a device to be configured as an edge node.

To specify an edge node:

Step	Command
1. Enter system view.	<b>system-view</b>
2. Enter RRPP domain view.	<b>rrpp domain</b> <i>domain-id</i>
3. Specify the current device as a transit node of the primary ring, and specify the primary port and the secondary port.	<b>ring</b> <i>ring-id</i> <b>node-mode transit</b> [ <b>primary-port</b> <i>interface-type interface-number</i> ] [ <b>secondary-port</b> <i>interface-type interface-number</i> ] <b>level</b> <i>level-value</i>
4. Specify the current device as the edge node of a subring, and specify the edge port.	<b>ring</b> <i>ring-id</i> <b>node-mode edge</b> [ <b>edge-port</b> <i>interface-type interface-number</i> ]

## Specifying an assistant-edge node

When configuring an assistant-edge node, you must first configure the primary ring before configuring the subrings.

Perform this configuration on a device to be configured as an assistant-edge node.

To specify an assistant-edge node:

Step	Command
1. Enter system view.	<b>system-view</b>
2. Enter RRPP domain view.	<b>rrpp domain</b> <i>domain-id</i>
3. Specify the current device as a transit node of the primary ring, and specify the primary port and the secondary port.	<b>ring</b> <i>ring-id</i> <b>node-mode transit</b> [ <b>primary-port</b> <i>interface-type interface-number</i> ] [ <b>secondary-port</b> <i>interface-type interface-number</i> ] <b>level</b> <i>level-value</i>
4. Specify the current device as the assistant-edge node of the subring, and specify an edge port.	<b>ring</b> <i>ring-id</i> <b>node-mode assistant-edge</b> [ <b>edge-port</b> <i>interface-type interface-number</i> ]

# Activating an RRPP domain

To activate an RRPP domain on the current device, enable the RRPP protocol and RRPP rings for the RRPP domain on the current device.

To prevent Hello packets of subrings from being looped on the primary ring, enable the primary ring on its master node before enabling the subrings on their separate master nodes. On an edge node or assistant-edge node, enable/disable the primary ring and subrings separately:

- Enable the primary ring of an RRPP domain before enabling the subrings of the RRPP domain.
- Disable the primary ring of an RRPP domain after disabling all subrings of the RRPP domain.

Perform this operation on all nodes in the RRPP domain.

To activate an RRPP domain:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Enable RRPP.	<b>rrpp enable</b>	Disabled by default.
3. Enter RRPP domain view.	<b>rrpp domain</b> <i>domain-id</i>	N/A
4. Enable the specified RRPP ring.	<b>ring</b> <i>ring-id</i> <b>enable</b>	Disabled by default.

## Configuring RRPP timers

Perform this configuration on the master node of an RRPP domain.

To configure RRPP timers:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Enter RRPP domain view.	<b>rrpp domain</b> <i>domain-id</i>	N/A
3. Configure the Hello timer and Fail timer for the RRPP domain.	<b>timer hello-timer</b> <i>hello-value</i> <b>fail-timer</b> <i>fail-value</i>	By default, the Hello timer value is 1 second, and the Fail timer value is 3 seconds.

### NOTE:

- The Fail timer value must be equal to or greater than three times the Hello timer value.
- To avoid temporary loops when the primary ring fails in a dual-homed-ring network, make sure that the difference between the Fail timer value on the master node of the subring and that on the master node of the primary ring is greater than twice the Hello timer value of the master node of the subring.

## Configuring an RRPP ring group

To reduce Edge-Hello traffic, adopt the RRPP ring group mechanism by assigning subrings with the same edge node/assistant-edge node to an RRPP ring group. An RRPP ring group must be configured on both the edge node and the assistant-edge node and can only be configured on these two types of nodes.

Perform this configuration on both the edge node and the assistant-edge node in an RRPP domain.

## Configuration restrictions and guidelines

- You can assign a subring to only one RRPP ring group. Make sure the RRPP ring group configured on the edge node and the RRPP ring group configured on the assistant-edge node contain the same subrings. Otherwise, the RRPP ring group cannot operate properly.
- Make sure the subrings in an RRPP ring group share the same edge node and assistant-edge node and that the edge node and the assistant edge node have the same SRPTs.
- Make sure a device plays the same role on the subrings in an RRPP ring group. The role can be the edge node or the assistant-edge node.
- Make sure the RRPP ring group on the edge node and the RRPP ring group on the assistant-edge node have the same configurations and activation status.

- Make sure that all subrings in an RRPP ring group have the same SRPTs. If the SRPTs of these subrings are configured or modified differently, the RRPP ring group cannot operate properly.

## Configuration procedure

To configure an RRPP ring group:

Step	Command
1. Enter system view.	<b>system-view</b>
2. Create an RRPP ring group and enter RRPP ring group view.	<b>rrpp ring-group</b> <i>ring-group-id</i>
3. Assign the specified subrings to the RRPP ring group.	<b>domain</b> <i>domain-id</i> <b>ring</b> <i>ring-id-list</i>

## Displaying and maintaining RRPP

Task	Command	Remarks
Display brief RRPP information.	<b>display rrpp brief</b> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]	Available in any view
Display RRPP group configuration information.	<b>display rrpp ring-group</b> [ <i>ring-group-id</i> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]	Available in any view
Display detailed RRPP information.	<b>display rrpp verbose domain</b> <i>domain-id</i> [ <b>ring</b> <i>ring-id</i> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]	Available in any view
Display RRPP statistics.	<b>display rrpp statistics domain</b> <i>domain-id</i> [ <b>ring</b> <i>ring-id</i> ] [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]	Available in any view
Clear RRPP statistics.	<b>reset rrpp statistics domain</b> <i>domain-id</i> [ <b>ring</b> <i>ring-id</i> ]	Available in user view

## RRPP configuration examples

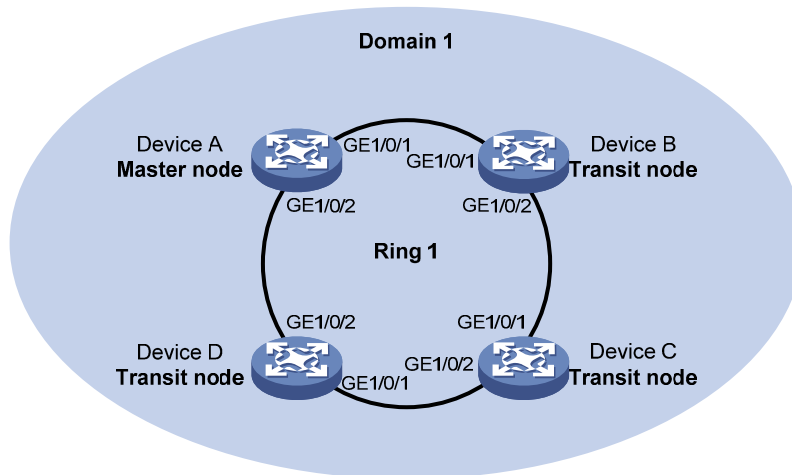
### Single ring configuration example

#### Networking requirements

As shown in [Figure 20](#),

- Device A, Device B, Device C, and Device D form RRPP domain 1. Specify the primary control VLAN of RRPP domain 1 as VLAN 4092. RRPP domain 1 protects VLANs 1 through 30.
- Device A, Device B, Device C, and Device D form primary ring 1.
- Specify Device A as the master node of primary ring 1, GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port.
- Specify Device B, Device C, and Device D as the transit nodes of primary ring 1. Specify their GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port.

Figure 20 Network diagram



## Configuration procedure

### 1. Configure Device A:

# Create VLANs 1 through 30, map these VLANs to MSTI 1, and activate the MST region configuration.

```
<DeviceA> system-view
[DeviceA] vlan 1 to 30
[DeviceA] stp region-configuration
[DeviceA-mst-region] instance 1 vlan 1 to 30
[DeviceA-mst-region] active region-configuration
[DeviceA-mst-region] quit
```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the two ports as trunk ports, and assign them to VLANs 1 through 30.

```
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] undo link-delay
[DeviceA-GigabitEthernet1/0/1] undo stp enable
[DeviceA-GigabitEthernet1/0/1] qos trust dot1p
[DeviceA-GigabitEthernet1/0/1] port link-type trunk
[DeviceA-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceA-GigabitEthernet1/0/1] quit
[DeviceA] interface gigabitethernet 1/0/2
[DeviceA-GigabitEthernet1/0/2] undo link-delay
[DeviceA-GigabitEthernet1/0/2] undo stp enable
[DeviceA-GigabitEthernet1/0/2] qos trust dot1p
[DeviceA-GigabitEthernet1/0/2] port link-type trunk
[DeviceA-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceA-GigabitEthernet1/0/2] quit
```

# Create RRPP domain 1. Configure VLAN 4092 as the primary control VLAN of RRPP domain 1, and configure the VLANs mapped to MSTI 1 as the protected VLANs of RRPP domain 1.

```
[DeviceA] rrpp domain 1
[DeviceA-rrpp-domain1] control-vlan 4092
```

```
[DeviceA-rrpp-domain1] protected-vlan reference-instance 1
Configure Device A as the master node of primary ring 1, with GigabitEthernet 1/0/1 as the
primary port and GigabitEthernet 1/0/2 as the secondary port, and enable ring 1.
[DeviceA-rrpp-domain1] ring 1 node-mode master primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 0
[DeviceA-rrpp-domain1] ring 1 enable
[DeviceA-rrpp-domain1] quit
```

# Enable RRPP.

```
[DeviceA] rrpp enable
```

## 2. Configure Device B:

# Create VLANs 1 through 30, map these VLANs to MSTI 1, and activate the MST region configuration.

```
<DeviceB> system-view
[DeviceB] vlan 1 to 30
[DeviceB] stp region-configuration
[DeviceB-mst-region] instance 1 vlan 1 to 30
[DeviceB-mst-region] active region-configuration
[DeviceB-mst-region] quit
```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the two ports as trunk ports, and assign them to VLANs 1 through 30.

```
[DeviceB] interface gigabitethernet 1/0/1
[DeviceB-GigabitEthernet1/0/1] undo link-delay
[DeviceB-GigabitEthernet1/0/1] undo stp enable
[DeviceB-GigabitEthernet1/0/1] qos trust dot1p
[DeviceB-GigabitEthernet1/0/1] port link-type trunk
[DeviceB-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceB-GigabitEthernet1/0/1] quit
[DeviceB] interface gigabitethernet 1/0/2
[DeviceB-GigabitEthernet1/0/2] undo link-delay
[DeviceB-GigabitEthernet1/0/2] undo stp enable
[DeviceB-GigabitEthernet1/0/2] qos trust dot1p
[DeviceB-GigabitEthernet1/0/2] port link-type trunk
[DeviceB-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceB-GigabitEthernet1/0/2] quit
```

# Create RRPP domain 1. Configure VLAN 4092 as the primary control VLAN of RRPP domain 1, and configure the VLANs mapped to MSTI 1 as the protected VLANs of RRPP domain 1.

```
[DeviceB] rrpp domain 1
[DeviceB-rrpp-domain1] control-vlan 4092
[DeviceB-rrpp-domain1] protected-vlan reference-instance 1
```

# Configure Device B as the transit node of primary ring 1, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port, and enable ring 1.

```
[DeviceB-rrpp-domain1] ring 1 node-mode transit primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 0
[DeviceB-rrpp-domain1] ring 1 enable
[DeviceB-rrpp-domain1] quit
```

# Enable RRPP.

```
[DeviceB] rrpp enable
```

3. Configure Device C:

The configuration on Device C is similar to that on Device B and is not shown here.

4. Configure Device D:

The configuration on Device D is similar to that on Device B and is not shown here.

5. Verify the configuration:

Use the **display** command to view RRPP configuration and operational information on each device.

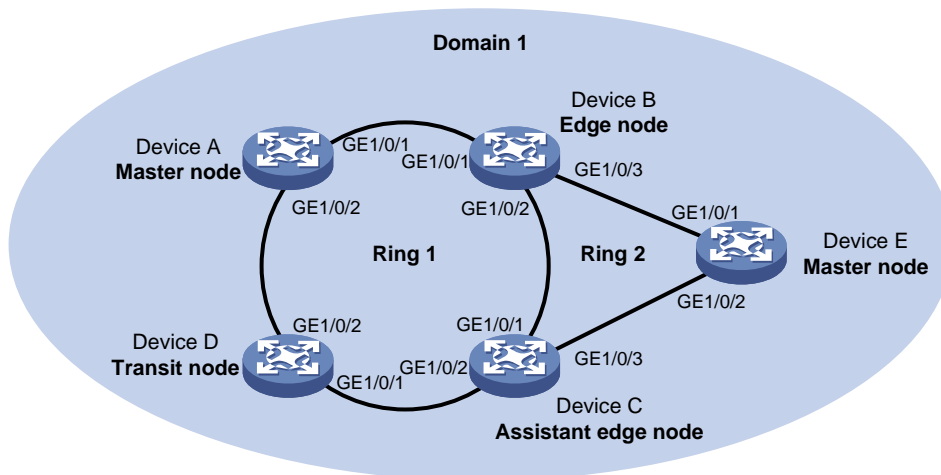
## Intersecting ring configuration example

### Networking requirements

As shown in [Figure 21](#),

- Device A, Device B, Device C, Device D, and Device E form RRPP domain 1. VLAN 4092 is the primary control VLAN of RRPP domain 1, and RRPP domain 1 protects VLANs 1 through 30.
- Device A, Device B, Device C, and Device D form primary ring 1, and Device B, Device C and Device E form subring 2.
- Device A is the master node of primary ring 1, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 the secondary port.
- Device E is the master node of subring 2, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 the secondary port.
- Device B is the transit node of primary ring 1 and the edge node of subring 2, and GigabitEthernet 1/0/3 is the edge port.
- Device C is the transit node of primary ring 1 and the assistant-edge node of subring 1, and GigabitEthernet 1/0/3 is the edge port.
- Device D is the transit node of primary ring 1, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 the secondary port.

**Figure 21 Network diagram**



### Configuration procedure

1. Configure Device A:

# Create VLANs 1 through 30, map these VLANs to MSTI 1, and activate the MST region configuration.

```
<DeviceA> system-view
[DeviceA] vlan 1 to 30
[DeviceA] stp region-configuration
[DeviceA-mst-region] instance 1 vlan 1 to 30
[DeviceA-mst-region] active region-configuration
[DeviceA-mst-region] quit
```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the two ports as trunk ports, and assign them to VLANs 1 through 30.

```
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] undo link-delay
[DeviceA-GigabitEthernet1/0/1] undo stp enable
[DeviceA-GigabitEthernet1/0/1] qos trust dot1p
[DeviceA-GigabitEthernet1/0/1] port link-type trunk
[DeviceA-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceA-GigabitEthernet1/0/1] quit
[DeviceA] interface gigabitethernet 1/0/2
[DeviceA-GigabitEthernet1/0/2] undo link-delay
[DeviceA-GigabitEthernet1/0/2] undo stp enable
[DeviceA-GigabitEthernet1/0/2] qos trust dot1p
[DeviceA-GigabitEthernet1/0/2] port link-type trunk
[DeviceA-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceA-GigabitEthernet1/0/2] quit
```

# Create RRPP domain 1. Configure VLAN 4092 as the primary control VLAN of RRPP domain 1, and configure the VLANs mapped to MSTI 1 as the protected VLANs of RRPP domain 1.

```
[DeviceA] rrpp domain 1
[DeviceA-rrpp-domain1] control-vlan 4092
[DeviceA-rrpp-domain1] protected-vlan reference-instance 1
```

# Configure Device A as the master node of primary ring 1, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port, and enable ring 1.

```
[DeviceA-rrpp-domain1] ring 1 node-mode master primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 0
[DeviceA-rrpp-domain1] ring 1 enable
[DeviceA-rrpp-domain1] quit
```

# Enable RRPP.

```
[DeviceA] rrpp enable
```

## 2. Configure Device B:

# Create VLANs 1 through 30, map these VLANs to MSTI 1, and activate the MST region configuration.

```
<DeviceB> system-view
[DeviceB] vlan 1 to 30
[DeviceB] stp region-configuration
[DeviceB-mst-region] instance 1 vlan 1 to 30
[DeviceB-mst-region] active region-configuration
[DeviceB-mst-region] quit
```



# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1, GigabitEthernet 1/0/2, and GigabitEthernet 1/0/3, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the three ports as trunk ports, and assign them to VLANs 1 through 30.

```
[DeviceB] interface gigabitethernet 1/0/1
[DeviceB-GigabitEthernet1/0/1] undo link-delay
[DeviceB-GigabitEthernet1/0/1] undo stp enable
[DeviceB-GigabitEthernet1/0/1] qos trust dot1p
[DeviceB-GigabitEthernet1/0/1] port link-type trunk
[DeviceB-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceB-GigabitEthernet1/0/1] quit
[DeviceB] interface gigabitethernet 1/0/2
[DeviceB-GigabitEthernet1/0/2] undo link-delay
[DeviceB-GigabitEthernet1/0/2] undo stp enable
[DeviceB-GigabitEthernet1/0/2] qos trust dot1p
[DeviceB-GigabitEthernet1/0/2] port link-type trunk
[DeviceB-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceB-GigabitEthernet1/0/2] quit
[DeviceB] interface gigabitethernet 1/0/3
[DeviceB-GigabitEthernet1/0/3] undo link-delay
[DeviceB-GigabitEthernet1/0/3] undo stp enable
[DeviceB-GigabitEthernet1/0/3] qos trust dot1p
[DeviceB-GigabitEthernet1/0/3] port link-type trunk
[DeviceB-GigabitEthernet1/0/3] port trunk permit vlan 1 to 30
[DeviceB-GigabitEthernet1/0/3] quit
```

# Create RRPP domain 1. Configure VLAN 4092 as the primary control VLAN of RRPP domain 1, and configure the VLANs mapped to MSTI 1 as the protected VLANs of RRPP domain 1.

```
[DeviceB] rrpp domain 1
[DeviceB-rrpp-domain1] control-vlan 4092
[DeviceB-rrpp-domain1] protected-vlan reference-instance 1
```

# Configure Device B as a transit node of primary ring 1, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port, and enable ring 1.

```
[DeviceB-rrpp-domain1] ring 1 node-mode transit primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 0
[DeviceB-rrpp-domain1] ring 1 enable
```

# Configure Device B as the edge node of subring 2, with GigabitEthernet 1/0/3 as the edge port, and enable ring 2.

```
[DeviceB-rrpp-domain1] ring 2 node-mode edge edge-port gigabitethernet 1/0/3
[DeviceB-rrpp-domain1] ring 2 enable
[DeviceB-rrpp-domain1] quit
```

# Enable RRPP.

```
[DeviceB] rrpp enable
```

### 3. Configure Device C:

# Create VLANs 1 through 30, map these VLANs to MSTI 1, and activate the MST region configuration.

```
<DeviceC> system-view
[DeviceC] vlan 1 to 30
[DeviceC] stp region-configuration
```

```
[DeviceC-mst-region] instance 1 vlan 1 to 30
[DeviceC-mst-region] active region-configuration
[DeviceC-mst-region] quit
```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1, GigabitEthernet 1/0/2, and GigabitEthernet 1/0/3, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the three ports as trunk ports, and assign them to VLANs 1 through 30.

```
[DeviceC] interface gigabitethernet 1/0/1
[DeviceC-GigabitEthernet1/0/1] undo link-delay
[DeviceC-GigabitEthernet1/0/1] undo stp enable
[DeviceC-GigabitEthernet1/0/1] qos trust dot1p
[DeviceC-GigabitEthernet1/0/1] port link-type trunk
[DeviceC-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceC-GigabitEthernet1/0/1] quit
[DeviceC] interface gigabitethernet 1/0/2
[DeviceC-GigabitEthernet1/0/2] undo link-delay
[DeviceC-GigabitEthernet1/0/2] undo stp enable
[DeviceC-GigabitEthernet1/0/2] qos trust dot1p
[DeviceC-GigabitEthernet1/0/2] port link-type trunk
[DeviceC-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceC-GigabitEthernet1/0/2] quit
[DeviceC] interface gigabitethernet 1/0/3
[DeviceC-GigabitEthernet1/0/3] undo link-delay
[DeviceC-GigabitEthernet1/0/3] undo stp enable
[DeviceC-GigabitEthernet1/0/3] qos trust dot1p
[DeviceC-GigabitEthernet1/0/3] port link-type trunk
[DeviceC-GigabitEthernet1/0/3] port trunk permit vlan 1 to 30
[DeviceC-GigabitEthernet1/0/3] quit
```

# Create RRPP domain 1. Configure VLAN 4092 as the primary control VLAN of RRPP domain 1, and configure VLANs mapped to MSTI 1 as the protected VLANs of RRPP domain 1.

```
[DeviceC] rrpp domain 1
[DeviceC-rrpp-domain1] control-vlan 4092
[DeviceC-rrpp-domain1] protected-vlan reference-instance 1
```

# Configure Device C as a transit node of primary ring 1, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port, and enable ring 1.

```
[DeviceC-rrpp-domain1] ring 1 node-mode transit primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 0
[DeviceC-rrpp-domain1] ring 1 enable
```

# Configure Device C as the assistant-edge node of subring 2, with GigabitEthernet 1/0/3 as the edge port, and enable ring 2.

```
[DeviceC-rrpp-domain1] ring 2 node-mode assistant-edge edge-port gigabitethernet
1/0/3
[DeviceC-rrpp-domain1] ring 2 enable
[DeviceC-rrpp-domain1] quit
```

# Enable RRPP.

```
[DeviceC] rrpp enable
```

#### 4. Configure Device D:

# Create VLANs 1 through 30, map these VLANs to MSTI 1, and activate the MST region configuration.

```
<DeviceD> system-view
[DeviceD] vlan 1 to 30
[DeviceD] stp region-configuration
[DeviceD-mst-region] instance 1 vlan 1 to 30
[DeviceD-mst-region] active region-configuration
[DeviceD-mst-region] quit
```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the two ports as trunk ports, and assign them to VLANs 1 through 30.

```
[DeviceD] interface gigabitethernet 1/0/1
[DeviceD-GigabitEthernet1/0/1] undo link-delay
[DeviceD-GigabitEthernet1/0/1] undo stp enable
[DeviceD-GigabitEthernet1/0/1] qos trust dot1p
[DeviceD-GigabitEthernet1/0/1] port link-type trunk
[DeviceD-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceD-GigabitEthernet1/0/1] quit
[DeviceD] interface gigabitethernet 1/0/2
[DeviceD-GigabitEthernet1/0/2] undo link-delay
[DeviceD-GigabitEthernet1/0/2] undo stp enable
[DeviceD-GigabitEthernet1/0/2] qos trust dot1p
[DeviceD-GigabitEthernet1/0/2] port link-type trunk
[DeviceD-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceD-GigabitEthernet1/0/2] quit
```

# Create RRPP domain 1. Configure VLAN 4092 as the primary control VLAN of RRPP domain 1, and configure VLANs mapped to MSTI 1 as the protected VLANs of RRPP domain 1.

```
[DeviceD] rrpp domain 1
[DeviceD-rrpp-domain1] control-vlan 4092
[DeviceD-rrpp-domain1] protected-vlan reference-instance 1
```

# Configure Device D as the transit node of primary ring 1, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port, and enable ring 1.

```
[DeviceD-rrpp-domain1] ring 1 node-mode transit primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 0
[DeviceD-rrpp-domain1] ring 1 enable
[DeviceD-rrpp-domain1] quit
```

# Enable RRPP.

```
[DeviceD] rrpp enable
```

## 5. Configure Device E:

# Create VLANs 1 through 30, map these VLANs to MSTI 1, and activate the MST region configuration.

```
<DeviceE> system-view
[DeviceE] vlan 1 to 30
[DeviceE] stp region-configuration
[DeviceE-mst-region] instance 1 vlan 1 to 30
[DeviceE-mst-region] active region-configuration
[DeviceE-mst-region] quit
```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the two ports as trunk ports, and assign them to VLANs 1 through 30.

```
[DeviceE] interface gigabitethernet 1/0/1
[DeviceE-GigabitEthernet1/0/1] undo link-delay
[DeviceE-GigabitEthernet1/0/1] undo stp enable
[DeviceE-GigabitEthernet1/0/1] qos trust dot1p
[DeviceE-GigabitEthernet1/0/1] port link-type trunk
[DeviceE-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceE-GigabitEthernet1/0/1] quit
[DeviceE] interface gigabitethernet 1/0/2
[DeviceE-GigabitEthernet1/0/2] undo link-delay
[DeviceE-GigabitEthernet1/0/2] undo stp enable
[DeviceE-GigabitEthernet1/0/2] qos trust dot1p
[DeviceE-GigabitEthernet1/0/2] port link-type trunk
[DeviceE-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceE-GigabitEthernet1/0/2] quit
```

# Create RRPP domain 1. Configure VLAN 4092 as the primary control VLAN of RRPP domain 1, and configure VLANs mapped to MSTI 1 as the protected VLANs of RRPP domain 1.

```
[DeviceE] rrpp domain 1
[DeviceE-rrpp-domain1] control-vlan 4092
[DeviceE-rrpp-domain1] protected-vlan reference-instance 1
```

# Configure Device E as the master node of subring 2, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port, and enable ring 2.

```
[DeviceE-rrpp-domain1] ring 2 node-mode master primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 1
[DeviceE-rrpp-domain1] ring 2 enable
[DeviceE-rrpp-domain1] quit
```

# Enable RRPP.

```
[DeviceE] rrpp enable
```

#### 6. Verify the configuration:

Use the **display** command to view RRPP configuration and operational information on each device.

## Dual homed rings configuration example

### Networking requirements

As shown in [Figure 22](#),

- Device A through Device H form RRPP domain 1. Specify the primary control VLAN of RRPP domain 1 as VLAN 4092, and specify that RRPP domain 1 protects VLANs 1 through 30.
- Device A through Device D form primary ring 1. Device A, Device B, and Device E form subring 2. Device A, Device B, and Device F form subring 3. Device C, Device D, and Device G form subring 4. Device C, Device D, and Device H form subring 5.
- Specify Device A as the master node of primary ring 1, GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port. Specify Device E as the master node of subring 2, GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port. Specify Device F as the master node of subring 3, GigabitEthernet 1/0/1 as the primary port and

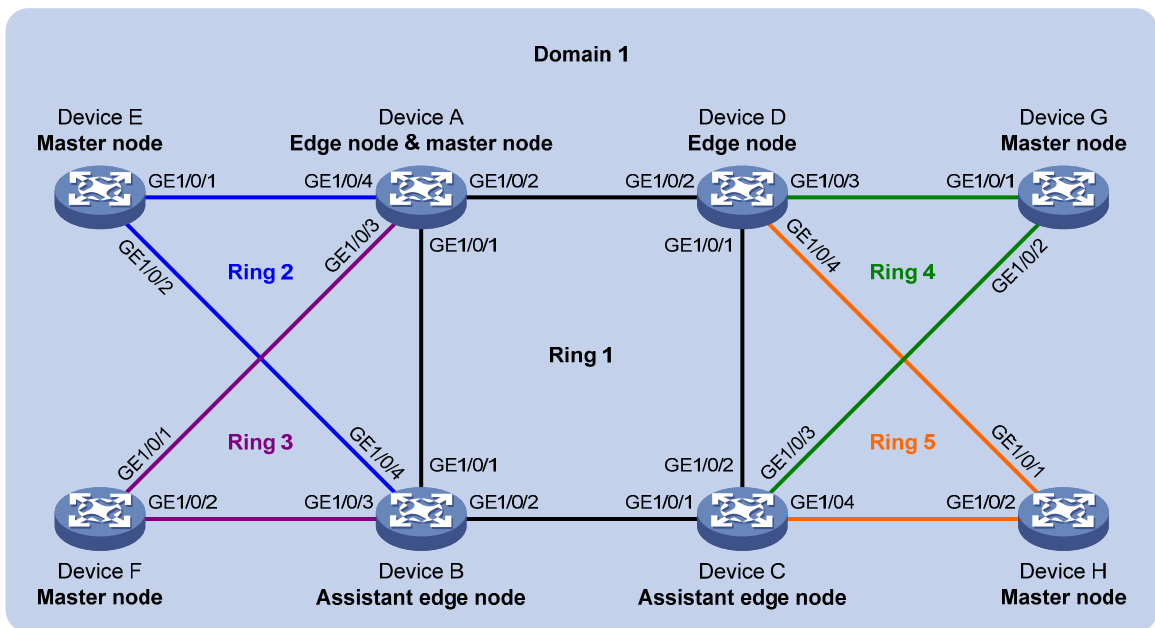
GigabitEthernet 1/0/2 as the secondary port. Specify Device G as the master node of subring 4, GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port. Specify Device H as the master node of subring 5, GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port.

- Specify Device A as the edge node of the connected subrings, its GigabitEthernet 1/0/3 and GigabitEthernet 1/0/4 as the edge ports. Specify Device D as the transit node of the primary ring and edge node of the connected subrings, its GigabitEthernet 1/0/3 and GigabitEthernet 1/0/4 as the edge ports. Specify Device B and Device C as the transit node of the primary ring and assistant-edge nodes of the connected subrings, their GigabitEthernet 1/0/3 and GigabitEthernet 1/0/4 as the edge ports.

**NOTE:**

Configure the primary and secondary ports on the master nodes properly to make sure that other protocols still work normally when data VLANs are denied by the secondary ports.

**Figure 22 Network diagram**



**Configuration procedure**

- Configure Device A:

# Create VLANs 1 through 30, map these VLANs to MSTI 1, and activate the MST region configuration.

```
<DeviceA> system-view
[DeviceA] vlan 1 to 30
[DeviceA] stp region-configuration
[DeviceA-mst-region] instance 1 vlan 1 to 30
[DeviceA-mst-region] active region-configuration
[DeviceA-mst-region] quit
```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 through GigabitEthernet 1/0/4, disable the spanning tree feature, and set the trusted packet priority type

to 802.1p priority. Configure the four ports as trunk ports, and assign them to VLANs 1 through 30.

```
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] undo link-delay
[DeviceA-GigabitEthernet1/0/1] undo stp enable
[DeviceA-GigabitEthernet1/0/1] qos trust dot1p
[DeviceA-GigabitEthernet1/0/1] port link-type trunk
[DeviceA-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceA-GigabitEthernet1/0/1] quit
[DeviceA] interface gigabitethernet 1/0/2
[DeviceA-GigabitEthernet1/0/2] undo link-delay
[DeviceA-GigabitEthernet1/0/2] undo stp enable
[DeviceA-GigabitEthernet1/0/2] qos trust dot1p
[DeviceA-GigabitEthernet1/0/2] port link-type trunk
[DeviceA-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceA-GigabitEthernet1/0/2] quit
[DeviceA] interface gigabitethernet 1/0/3
[DeviceA-GigabitEthernet1/0/3] undo link-delay
[DeviceA-GigabitEthernet1/0/3] undo stp enable
[DeviceA-GigabitEthernet1/0/3] qos trust dot1p
[DeviceA-GigabitEthernet1/0/3] port link-type trunk
[DeviceA-GigabitEthernet1/0/3] port trunk permit vlan 1 to 30
[DeviceA-GigabitEthernet1/0/3] quit
[DeviceA] interface gigabitethernet 1/0/4
[DeviceA-GigabitEthernet1/0/4] undo link-delay
[DeviceA-GigabitEthernet1/0/4] undo stp enable
[DeviceA-GigabitEthernet1/0/4] qos trust dot1p
[DeviceA-GigabitEthernet1/0/4] port link-type trunk
[DeviceA-GigabitEthernet1/0/4] port trunk permit vlan 1 to 30
[DeviceA-GigabitEthernet1/0/4] quit
```

**# Create RRPP domain 1. Configure VLAN 4092 as the primary control VLAN of RRPP domain 1, and configure the VLANs mapped to MSTI 1 as the protected VLANs of RRPP domain 1.**

```
[DeviceA] rrpp domain 1
[DeviceA-rrpp-domain1] control-vlan 4092
[DeviceA-rrpp-domain1] protected-vlan reference-instance 1
```

**# Configure Device A as the master node of primary ring 1, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port, and enable ring 1.**

```
[DeviceA-rrpp-domain1] ring 1 node-mode master primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 0
[DeviceA-rrpp-domain1] ring 1 enable
```

**# Configure Device A as the edge node of subring 2, with GigabitEthernet 1/0/4 as the edge port, and enable subring 2.**

```
[DeviceA-rrpp-domain1] ring 2 node-mode edge edge-port gigabitethernet 1/0/4
[DeviceA-rrpp-domain1] ring 2 enable
```

**# Configure Device A as the edge node of subring 3, with GigabitEthernet 1/0/3 as the edge port, and enable subring 3.**

```
[DeviceA-rrpp-domain1] ring 3 node-mode edge edge-port gigabitethernet 1/0/3
[DeviceA-rrpp-domain1] ring 3 enable
```

```
[DeviceA-rrpp-domain1] quit
```

```
Enable RRPP.
```

```
[DeviceA] rrpp enable
```

## 2. Configure Device B:

```
Create VLANs 1 through 30, map these VLANs to MSTI 1, and activate the MST region configuration.
```

```
<DeviceB> system-view
```

```
[DeviceB] vlan 1 to 30
```

```
[DeviceB] stp region-configuration
```

```
[DeviceB-mst-region] instance 1 vlan 1 to 30
```

```
[DeviceB-mst-region] active region-configuration
```

```
[DeviceB-mst-region] quit
```

```
Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 through GigabitEthernet 1/0/4, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the four ports as trunk ports, and assign them to VLANs 1 through 30.
```

```
[DeviceB] interface gigabitethernet 1/0/1
```

```
[DeviceB-GigabitEthernet1/0/1] undo link-delay
```

```
[DeviceB-GigabitEthernet1/0/1] undo stp enable
```

```
[DeviceB-GigabitEthernet1/0/1] qos trust dot1p
```

```
[DeviceB-GigabitEthernet1/0/1] port link-type trunk
```

```
[DeviceB-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
```

```
[DeviceB-GigabitEthernet1/0/1] quit
```

```
[DeviceB] interface gigabitethernet 1/0/2
```

```
[DeviceB-GigabitEthernet1/0/2] undo link-delay
```

```
[DeviceB-GigabitEthernet1/0/2] undo stp enable
```

```
[DeviceB-GigabitEthernet1/0/2] qos trust dot1p
```

```
[DeviceB-GigabitEthernet1/0/2] port link-type trunk
```

```
[DeviceB-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
```

```
[DeviceB-GigabitEthernet1/0/2] quit
```

```
[DeviceB] interface gigabitethernet 1/0/3
```

```
[DeviceB-GigabitEthernet1/0/3] undo link-delay
```

```
[DeviceB-GigabitEthernet1/0/3] undo stp enable
```

```
[DeviceB-GigabitEthernet1/0/3] qos trust dot1p
```

```
[DeviceB-GigabitEthernet1/0/3] port link-type trunk
```

```
[DeviceB-GigabitEthernet1/0/3] port trunk permit vlan 1 to 30
```

```
[DeviceB-GigabitEthernet1/0/3] quit
```

```
[DeviceB] interface gigabitethernet 1/0/4
```

```
[DeviceB-GigabitEthernet1/0/4] undo link-delay
```

```
[DeviceB-GigabitEthernet1/0/4] undo stp enable
```

```
[DeviceB-GigabitEthernet1/0/4] qos trust dot1p
```

```
[DeviceB-GigabitEthernet1/0/4] port link-type trunk
```

```
[DeviceB-GigabitEthernet1/0/4] port trunk permit vlan 1 to 30
```

```
[DeviceB-GigabitEthernet1/0/4] quit
```

```
Create RRPP domain 1. Configure VLAN 4092 as the primary control VLAN of RRPP domain 1, and configure the VLANs mapped to MSTI 1 as the protected VLANs of RRPP domain 1.
```

```
[DeviceB] rrpp domain 1
```

```
[DeviceB-rrpp-domain1] control-vlan 4092
```

```

[DeviceB-rrpp-domain1] protected-vlan reference-instance 1
Configure Device B as the transit node of primary ring 1, with GigabitEthernet 1/0/1 as the
primary port and GigabitEthernet 1/0/2 as the secondary port, and enable ring 1.
[DeviceB-rrpp-domain1] ring 1 node-mode transit primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 0
[DeviceB-rrpp-domain1] ring 1 enable
Configure Device B as the assistant-edge node of subring 2, with GigabitEthernet 1/0/4 as the
edge port, and enable subring 2.
[DeviceB-rrpp-domain1] ring 2 node-mode assistant-edge edge-port gigabitethernet
1/0/4
[DeviceB-rrpp-domain1] ring 2 enable
Configure Device B as the assistant-edge node of subring 3, with GigabitEthernet 1/0/3 as the
edge port, and enable subring 3.
[DeviceB-rrpp-domain1] ring 3 node-mode assistant-edge edge-port gigabitethernet
1/0/3
[DeviceB-rrpp-domain1] ring 3 enable
[DeviceB-rrpp-domain1] quit
Enable RRPP.
[DeviceB] rrpp enable

```

### 3. Configure Device C:

```

Create VLANs 1 through 30, map these VLANs to MSTI 1, and activate the MST region
configuration.
<DeviceC> system-view
[DeviceC] vlan 1 to 30
[DeviceC] stp region-configuration
[DeviceC-mst-region] instance 1 vlan 1 to 30
[DeviceC-mst-region] active region-configuration
[DeviceC-mst-region] quit
Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 through
GigabitEthernet 1/0/4, disable the spanning tree feature, and set the trusted packet priority type
to 802.1p priority. Configure the four ports as trunk ports, and assign them to VLANs 1 through
30.
[DeviceC] interface gigabitethernet 1/0/1
[DeviceC-GigabitEthernet1/0/1] undo link-delay
[DeviceC-GigabitEthernet1/0/1] undo stp enable
[DeviceC-GigabitEthernet1/0/1] qos trust dot1p
[DeviceC-GigabitEthernet1/0/1] port link-type trunk
[DeviceC-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceC-GigabitEthernet1/0/1] quit
[DeviceC] interface gigabitethernet 1/0/2
[DeviceC-GigabitEthernet1/0/2] undo link-delay
[DeviceC-GigabitEthernet1/0/2] undo stp enable
[DeviceC-GigabitEthernet1/0/2] qos trust dot1p
[DeviceC-GigabitEthernet1/0/2] port link-type trunk
[DeviceC-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceC-GigabitEthernet1/0/2] quit
[DeviceC] interface gigabitethernet 1/0/3
[DeviceC-GigabitEthernet1/0/3] undo link-delay

```



```

[DeviceC-GigabitEthernet1/0/3] undo stp enable
[DeviceC-GigabitEthernet1/0/3] qos trust dot1p
[DeviceC-GigabitEthernet1/0/3] port link-type trunk
[DeviceC-GigabitEthernet1/0/3] port trunk permit vlan 1 to 30
[DeviceC-GigabitEthernet1/0/3] quit
[DeviceC] interface gigabitethernet 1/0/4
[DeviceC-GigabitEthernet1/0/4] undo link-delay
[DeviceC-GigabitEthernet1/0/4] undo stp enable
[DeviceC-GigabitEthernet1/0/4] qos trust dot1p
[DeviceC-GigabitEthernet1/0/4] port link-type trunk
[DeviceC-GigabitEthernet1/0/4] port trunk permit vlan 1 to 30
[DeviceC-GigabitEthernet1/0/4] quit
Create RRPP domain 1. Configure VLAN 4092 as the primary control VLAN of RRPP domain 1,
and configure the VLANs mapped to MSTI 1 as the protected VLANs of RRPP domain 1.
[DeviceC] rrpp domain 1
[DeviceC-rrpp-domain1] control-vlan 4092
[DeviceC-rrpp-domain1] protected-vlan reference-instance 1
Configure Device C as the transit node of primary ring 1, with GigabitEthernet 1/0/1 as the
primary port and GigabitEthernet 1/0/2 as the secondary port, and enable ring 1.
[DeviceC-rrpp-domain1] ring 1 node-mode transit primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 0
[DeviceC-rrpp-domain1] ring 1 enable
Configure Device C as the assistant-edge node of subring 4, with GigabitEthernet 1/0/3 as the
edge port, and enable subring 4.
[DeviceC-rrpp-domain1] ring 4 node-mode assistant-edge edge-port gigabitethernet
1/0/3
[DeviceC-rrpp-domain1] ring 4 enable
Configure Device C as the assistant-edge node of subring 5, with GigabitEthernet 1/0/4 as the
edge port, and enable subring 5.
[DeviceC-rrpp-domain1] ring 5 node-mode assistant-edge edge-port gigabitethernet
1/0/4
[DeviceC-rrpp-domain1] ring 5 enable
[DeviceC-rrpp-domain1] quit
Enable RRPP.
[DeviceC] rrpp enable

```

#### 4. Configure Device D:

# Create VLANs 1 through 30, map these VLANs to MSTI 1, and activate the MST region configuration.

```

<DeviceD> system-view
[DeviceD] vlan 1 to 30
[DeviceD] stp region-configuration
[DeviceD-mst-region] instance 1 vlan 1 to 30
[DeviceD-mst-region] active region-configuration
[DeviceD-mst-region] quit

```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 through GigabitEthernet 1/0/4, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the four ports as trunk ports, and assign them to VLANs 1 through 30.

```

[DeviceD] interface gigabitethernet 1/0/1
[DeviceD-GigabitEthernet1/0/1] undo link-delay
[DeviceD-GigabitEthernet1/0/1] undo stp enable
[DeviceD-GigabitEthernet1/0/1] qos trust dot1p
[DeviceD-GigabitEthernet1/0/1] port link-type trunk
[DeviceD-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceD-GigabitEthernet1/0/1] quit
[DeviceD] interface gigabitethernet 1/0/2
[DeviceD-GigabitEthernet1/0/2] undo link-delay
[DeviceD-GigabitEthernet1/0/2] undo stp enable
[DeviceD-GigabitEthernet1/0/2] qos trust dot1p
[DeviceD-GigabitEthernet1/0/2] port link-type trunk
[DeviceD-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceD-GigabitEthernet1/0/2] quit
[DeviceD] interface gigabitethernet 1/0/3
[DeviceD-GigabitEthernet1/0/3] undo link-delay
[DeviceD-GigabitEthernet1/0/3] undo stp enable
[DeviceD-GigabitEthernet1/0/3] qos trust dot1p
[DeviceD-GigabitEthernet1/0/3] port link-type trunk
[DeviceD-GigabitEthernet1/0/3] port trunk permit vlan 1 to 30
[DeviceD-GigabitEthernet1/0/3] quit
[DeviceD] interface gigabitethernet 1/0/4
[DeviceD-GigabitEthernet1/0/4] undo link-delay
[DeviceD-GigabitEthernet1/0/4] undo stp enable
[DeviceD-GigabitEthernet1/0/4] qos trust dot1p
[DeviceD-GigabitEthernet1/0/4] port link-type trunk
[DeviceD-GigabitEthernet1/0/4] port trunk permit vlan 1 to 30
[DeviceD-GigabitEthernet1/0/4] quit
Create RRPP domain 1. Configure VLAN 4092 as the primary control VLAN of RRPP domain 1,
and configure the VLANs mapped to MSTI 1 as the protected VLANs of RRPP domain 1.
[DeviceD] rrpp domain 1
[DeviceD-rrpp-domain1] control-vlan 4092
[DeviceD-rrpp-domain1] protected-vlan reference-instance 1
Configure Device D as the transit node of primary ring 1, with GigabitEthernet 1/0/1 as the
primary port and GigabitEthernet 1/0/2 as the secondary port, and enable ring 1.
[DeviceD-rrpp-domain1] ring 1 node-mode transit primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 0
[DeviceD-rrpp-domain1] ring 1 enable
Configure Device D as the edge node of subring 4, with GigabitEthernet 1/0/3 as the edge port,
and enable subring 4.
[DeviceD-rrpp-domain1] ring 4 node-mode edge edge-port gigabitethernet 1/0/3
[DeviceD-rrpp-domain1] ring 4 enable
Configure Device D as the edge node of subring 5, with GigabitEthernet 1/0/4 as the edge port,
and enable subring 5.
[DeviceD-rrpp-domain1] ring 5 node-mode edge edge-port gigabitethernet 1/0/4
[DeviceD-rrpp-domain1] ring 5 enable
[DeviceD-rrpp-domain1] quit
Enable RRPP.

```

```
[DeviceD] rrpp enable
```

## 5. Configure Device E:

# Create VLANs 1 through 30, map these VLANs to MSTI 1, and activate the MST region configuration.

```
<DeviceE> system-view
[DeviceE] vlan 1 to 30
[DeviceE] stp region-configuration
[DeviceE-mst-region] instance 1 vlan 1 to 30
[DeviceE-mst-region] active region-configuration
[DeviceE-mst-region] quit
```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the two ports as trunk ports, and assign them to VLANs 1 through 30.

```
[DeviceE] interface gigabitethernet 1/0/1
[DeviceE-GigabitEthernet1/0/1] undo link-delay
[DeviceE-GigabitEthernet1/0/1] undo stp enable
[DeviceE-GigabitEthernet1/0/1] qos trust dot1p
[DeviceE-GigabitEthernet1/0/1] port link-type trunk
[DeviceE-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceE-GigabitEthernet1/0/1] quit
[DeviceE] interface gigabitethernet 1/0/2
[DeviceE-GigabitEthernet1/0/2] undo link-delay
[DeviceE-GigabitEthernet1/0/2] undo stp enable
[DeviceE-GigabitEthernet1/0/2] qos trust dot1p
[DeviceE-GigabitEthernet1/0/2] port link-type trunk
[DeviceE-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceE-GigabitEthernet1/0/2] quit
```

# Create RRPP domain 1. Configure VLAN 4092 as the primary control VLAN of RRPP domain 1, and configure the VLANs mapped to MSTI 1 as the protected VLANs of RRPP domain 1.

```
[DeviceE] rrpp domain 1
[DeviceE-rrpp-domain1] control-vlan 4092
[DeviceE-rrpp-domain1] protected-vlan reference-instance 1
```

# Configure Device E as the master node of subring 2, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port, and enable subring 2.

```
[DeviceE-rrpp-domain1] ring 2 node-mode master primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 1
[DeviceE-rrpp-domain1] ring 2 enable
[DeviceE-rrpp-domain1] quit
```

# Enable RRPP.

```
[DeviceE] rrpp enable
```

## 6. Configure Device F:

# Create VLANs 1 through 30, map these VLANs to MSTI 1, and activate the MST region configuration.

```
<DeviceF> system-view
[DeviceF] vlan 1 to 30
[DeviceF] stp region-configuration
```

```
[DeviceF-mst-region] instance 1 vlan 1 to 30
[DeviceF-mst-region] active region-configuration
[DeviceF-mst-region] quit
```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the two ports as trunk ports, and assign them to VLANs 1 through 30.

```
[DeviceF] interface gigabitethernet 1/0/1
[DeviceF-GigabitEthernet1/0/1] undo link-delay
[DeviceF-GigabitEthernet1/0/1] undo stp enable
[DeviceF-GigabitEthernet1/0/1] qos trust dot1p
[DeviceF-GigabitEthernet1/0/1] port link-type trunk
[DeviceF-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceF-GigabitEthernet1/0/1] quit
[DeviceF] interface gigabitethernet 1/0/2
[DeviceF-GigabitEthernet1/0/2] undo link-delay
[DeviceF-GigabitEthernet1/0/2] undo stp enable
[DeviceF-GigabitEthernet1/0/2] qos trust dot1p
[DeviceF-GigabitEthernet1/0/2] port link-type trunk
[DeviceF-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceF-GigabitEthernet1/0/2] quit
```

# Create RRPP domain 1. Configure VLAN 4092 as the primary control VLAN of RRPP domain 1, and configure the VLANs mapped to MSTI 1 as the protected VLANs of RRPP domain 1.

```
[DeviceF] rrpp domain 1
[DeviceF-rrpp-domain1] control-vlan 4092
[DeviceF-rrpp-domain1] protected-vlan reference-instance 1
```

# Configure Device F as the master node of subring 3, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port, and enable subring 3.

```
[DeviceF-rrpp-domain1] ring 3 node-mode master primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 1
[DeviceF-rrpp-domain1] ring 3 enable
[DeviceF-rrpp-domain1] quit
```

# Enable RRPP.

```
[DeviceF] rrpp enable
```

## 7. Configure Device G:

# Create VLANs 1 through 30, map these VLANs to MSTI 1, and activate the MST region configuration.

```
<DeviceG> system-view
[DeviceG] vlan 1 to 30
[DeviceG] stp region-configuration
[DeviceG-mst-region] instance 1 vlan 1 to 30
[DeviceG-mst-region] active region-configuration
[DeviceG-mst-region] quit
```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the two ports as trunk ports, and assign them to VLANs 1 through 30.

```
[DeviceG] interface gigabitethernet 1/0/1
```

```

[DeviceG-GigabitEthernet1/0/1] undo link-delay
[DeviceG-GigabitEthernet1/0/1] undo stp enable
[DeviceG-GigabitEthernet1/0/1] qos trust dot1p
[DeviceG-GigabitEthernet1/0/1] port link-type trunk
[DeviceG-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceG-GigabitEthernet1/0/1] quit
[DeviceG] interface gigabitethernet 1/0/2
[DeviceG-GigabitEthernet1/0/2] undo link-delay
[DeviceG-GigabitEthernet1/0/2] undo stp enable
[DeviceG-GigabitEthernet1/0/2] qos trust dot1p
[DeviceG-GigabitEthernet1/0/2] port link-type trunk
[DeviceG-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceG-GigabitEthernet1/0/2] quit

```

# Create RRPP domain 1. Configure VLAN 4092 as the primary control VLAN of RRPP domain 1, and configure the VLANs mapped to MSTI 1 as the protected VLANs of RRPP domain 1.

```

[DeviceG] rrpp domain 1
[DeviceG-rrpp-domain1] control-vlan 4092
[DeviceG-rrpp-domain1] protected-vlan reference-instance 1

```

# Configure Device G as the master node of subring 4, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port, and enable subring 4.

```

[DeviceG-rrpp-domain1] ring 4 node-mode master primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 1
[DeviceG-rrpp-domain1] ring 4 enable
[DeviceG-rrpp-domain1] quit

```

# Enable RRPP.

```

[DeviceG] rrpp enable

```

## 8. Configure Device H:

# Create VLANs 1 through 30, map these VLANs to MSTI 1, and activate the MST region configuration.

```

<DeviceH> system-view
[DeviceH] vlan 1 to 30
[DeviceH] stp region-configuration
[DeviceH-mst-region] instance 1 vlan 1 to 30
[DeviceH-mst-region] active region-configuration
[DeviceH-mst-region] quit

```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the two ports as trunk ports, and assign them to VLANs 1 through 30.

```

[DeviceH] interface gigabitethernet 1/0/1
[DeviceH-GigabitEthernet1/0/1] undo link-delay
[DeviceH-GigabitEthernet1/0/1] undo stp enable
[DeviceH-GigabitEthernet1/0/1] qos trust dot1p
[DeviceH-GigabitEthernet1/0/1] port link-type trunk
[DeviceH-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceH-GigabitEthernet1/0/1] quit
[DeviceH] interface gigabitethernet 1/0/2
[DeviceH-GigabitEthernet1/0/2] undo link-delay

```

```
[DeviceH-GigabitEthernet1/0/2] undo stp enable
[DeviceH-GigabitEthernet1/0/2] qos trust dot1p
[DeviceH-GigabitEthernet1/0/2] port link-type trunk
[DeviceH-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceH-GigabitEthernet1/0/2] quit
```

# Create RRPP domain 1. Configure VLAN 4092 as the primary control VLAN of RRPP domain 1, and configure the VLANs mapped to MSTI 1 as the protected VLANs of RRPP domain 1.

```
[DeviceH] rrpp domain 1
[DeviceH-rrpp-domain1] control-vlan 4092
[DeviceH-rrpp-domain1] protected-vlan reference-instance 1
```

# Configure Device H as the master node of subring 5, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port, and enable subring 5.

```
[DeviceH-rrpp-domain1] ring 5 node-mode master primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 1
[DeviceH-rrpp-domain1] ring 5 enable
[DeviceH-rrpp-domain1] quit
```

# Enable RRPP.

```
[DeviceH] rrpp enable
```

#### 9. Verify the configuration:

Use the **display** command to view RRPP configuration and operational information on each device.

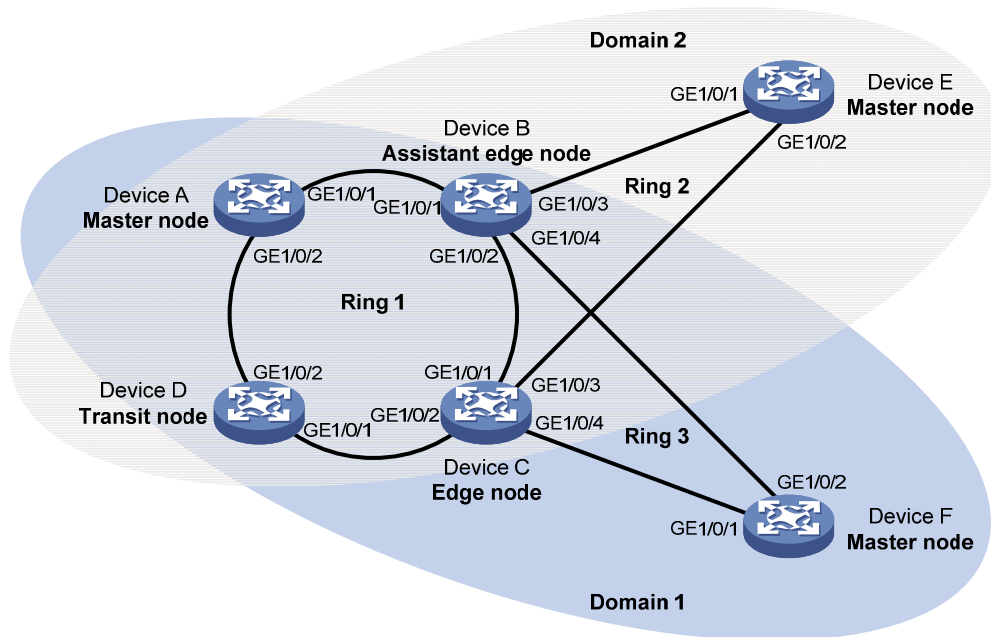
## Intersecting-ring load balancing configuration example

### Networking requirements

As shown in [Figure 23](#),

- Device A, Device B, Device C, Device D, and Device F form RRPP domain 1, and VLAN 100 is the primary control VLAN of the RRPP domain. Device A is the master node of the primary ring, Ring 1; Device D is the transit node of Ring 1; Device F is the master node of the subring Ring 3; Device C is the edge node of the subring Ring 3; Device B is the assistant-edge node of the subring Ring 3.
- Device A, Device B, Device C, Device D, and Device E form RRPP domain 2, and VLAN 105 is the primary control VLAN of the RRPP domain. Device A is the master node of the primary ring, Ring 1; Device D is the transit node of Ring 1; Device E is the master node of the subring Ring 2; Device C is the edge node of the subring Ring 2; Device B is the assistant-edge node of the subring Ring 2.
- Specify VLAN 1 as the protected VLAN of domain 1 and VLAN 2 the protected VLAN of domain 2. You can implement VLAN-based load balancing on Ring 1.
- Because the edge node and assistant-edge node of Ring 2 are the same as those of Ring 3 and the two subrings have the same SRPTs, you can add Ring 2 and Ring 3 to the RRPP ring group to reduce Edge-Hello traffic.

Figure 23 Network diagram



## Configuration procedure

### 1. Configure Device A:

# Create VLANs 1 and 2, map VLAN 1 to MSTI 1 and VLAN 2 to MSTI 2, and activate MST region configuration.

```
<DeviceA> system-view
[DeviceA] vlan 1 to 2
[DeviceA] stp region-configuration
[DeviceA-mst-region] instance 1 vlan 1
[DeviceA-mst-region] instance 2 vlan 2
[DeviceA-mst-region] active region-configuration
[DeviceA-mst-region] quit
```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the two ports as trunk ports, and assign them to VLAN 1 and VLAN 2.

```
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] undo link-delay
[DeviceA-GigabitEthernet1/0/1] undo stp enable
[DeviceA-GigabitEthernet1/0/1] qos trust dot1p
[DeviceA-GigabitEthernet1/0/1] port link-type trunk
[DeviceA-GigabitEthernet1/0/1] port trunk permit vlan 1 2
[DeviceA-GigabitEthernet1/0/1] quit
[DeviceA] interface gigabitethernet 1/0/2
[DeviceA-GigabitEthernet1/0/2] undo link-delay
[DeviceA-GigabitEthernet1/0/2] undo stp enable
[DeviceA-GigabitEthernet1/0/2] qos trust dot1p
[DeviceA-GigabitEthernet1/0/2] port link-type trunk
[DeviceA-GigabitEthernet1/0/2] port trunk permit vlan 1 2
```

```

[DeviceA-GigabitEthernet1/0/2] quit
Create RRPP domain 1. Configure VLAN 100 as the primary control VLAN of RRPP domain 1,
and configure the VLAN mapped to MSTI 1 as the protected VLAN of RRPP domain 1.
[DeviceA] rrpp domain 1
[DeviceA-rrpp-domain1] control-vlan 100
[DeviceA-rrpp-domain1] protected-vlan reference-instance 1
Configure Device A as the master node of primary ring 1, with GigabitEthernet 1/0/1 as the
primary port and GigabitEthernet 1/0/2 as the secondary port, and enable ring 1.
[DeviceA-rrpp-domain1] ring 1 node-mode master primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 0
[DeviceA-rrpp-domain1] ring 1 enable
[DeviceA-rrpp-domain1] quit
Create RRPP domain 2, configure VLAN 105 as the primary control VLAN of RRPP domain 2,
and configure the VLAN mapped to MSTI 2 as the protected VLAN of RRPP domain 2.
[DeviceA] rrpp domain 2
[DeviceA-rrpp-domain2] control-vlan 105
[DeviceA-rrpp-domain2] protected-vlan reference-instance 2
Configure Device A as the master node of primary ring 1, with GigabitEthernet 1/0/2 as the
master port and GigabitEthernet 1/0/1 as the secondary port, and enable ring 1.
[DeviceA-rrpp-domain2] ring 1 node-mode master primary-port gigabitethernet 1/0/2
secondary-port gigabitethernet 1/0/1 level 0
[DeviceA-rrpp-domain2] ring 1 enable
[DeviceA-rrpp-domain2] quit
Enable RRPP.
[DeviceA] rrpp enable

```

## 2. Configure Device B:

```

Create VLANs 1 and 2, map VLAN 1 to MSTI 1 and VLAN 2 to MSTI 2, and activate MST region
configuration.
<DeviceB> system-view
[DeviceB] vlan 1 to 2
[DeviceB] stp region-configuration
[DeviceB-mst-region] instance 1 vlan 1
[DeviceB-mst-region] instance 2 vlan 2
[DeviceB-mst-region] active region-configuration
[DeviceB-mst-region] quit
Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 and
GigabitEthernet 1/0/2, disable the spanning tree feature, and set the trusted packet priority type
to 802.1p priority. Configure the two ports as trunk ports, and assign them to VLAN 1 and VLAN
2.
[DeviceB] interface gigabitethernet 1/0/1
[DeviceB-GigabitEthernet1/0/1] undo link-delay
[DeviceB-GigabitEthernet1/0/1] undo stp enable
[DeviceB-GigabitEthernet1/0/1] qos trust dot1p
[DeviceB-GigabitEthernet1/0/1] port link-type trunk
[DeviceB-GigabitEthernet1/0/1] port trunk permit vlan 1 2
[DeviceB-GigabitEthernet1/0/1] quit
[DeviceB] interface gigabitethernet 1/0/2

```



```
[DeviceB-GigabitEthernet1/0/2] undo link-delay
[DeviceB-GigabitEthernet1/0/2] undo stp enable
[DeviceB-GigabitEthernet1/0/2] qos trust dot1p
[DeviceB-GigabitEthernet1/0/2] port link-type trunk
[DeviceB-GigabitEthernet1/0/2] port trunk permit vlan 1 2
[DeviceB-GigabitEthernet1/0/2] quit
```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/3, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the port as a trunk port, and assign it to VLAN 2.

```
[DeviceB] interface gigabitethernet 1/0/3
[DeviceB-GigabitEthernet1/0/3] undo link-delay
[DeviceB-GigabitEthernet1/0/3] undo stp enable
[DeviceB-GigabitEthernet1/0/3] qos trust dot1p
[DeviceB-GigabitEthernet1/0/3] port link-type trunk
[DeviceB-GigabitEthernet1/0/3] port trunk permit vlan 2
[DeviceB-GigabitEthernet1/0/3] quit
```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/4, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the port as a trunk port, and assign it to VLAN 1.

```
[DeviceB] interface gigabitethernet 1/0/4
[DeviceB-GigabitEthernet1/0/4] undo link-delay
[DeviceB-GigabitEthernet1/0/4] undo stp enable
[DeviceB-GigabitEthernet1/0/4] qos trust dot1p
[DeviceB-GigabitEthernet1/0/4] port link-type trunk
[DeviceB-GigabitEthernet1/0/4] port trunk permit vlan 1
[DeviceB-GigabitEthernet1/0/4] quit
```

# Create RRPP domain 1. Configure VLAN 100 as the primary control VLAN of RRPP domain 1, and configure the VLAN mapped to MSTI 1 as the protected VLAN of RRPP domain 1.

```
[DeviceB] rrpp domain 1
[DeviceB-rrpp-domain1] control-vlan 100
[DeviceB-rrpp-domain1] protected-vlan reference-instance 1
```

# Configure Device B as a transit node of primary ring 1 in RRPP domain 1, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port, and enable ring 1.

```
[DeviceB-rrpp-domain1] ring 1 node-mode transit primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 0
[DeviceB-rrpp-domain1] ring 1 enable
```

# Configure Device B as the assistant-edge node of subring 3 in RRPP domain 1, with GigabitEthernet 1/0/4 as the edge port, and enable subring 3.

```
[DeviceB-rrpp-domain1] ring 3 node-mode assistant-edge edge-port gigabitethernet
1/0/4
[DeviceB-rrpp-domain1] ring 3 enable
[DeviceB-rrpp-domain1] quit
```

# Create RRPP domain 2. Configure VLAN 105 as the primary control VLAN of RRPP domain 2, and configure the VLAN mapped to MSTI 2 as the protected VLAN of RRPP domain 2.

```
[DeviceB] rrpp domain 2
[DeviceB-rrpp-domain2] control-vlan 105
[DeviceB-rrpp-domain2] protected-vlan reference-instance 2
```

# Configure Device B as the transit node of primary ring 1, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port, and enable ring 1.

```
[DeviceB-rrpp-domain2] ring 1 node-mode transit primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 0
```

```
[DeviceB-rrpp-domain2] ring 1 enable
```

# Configure Device B as the assistant-edge node of subring 2 in RRPP domain 2, with GigabitEthernet 1/0/3 as the edge port, and enable subring 2.

```
[DeviceB-rrpp-domain2] ring 2 node-mode assistant-edge edge-port gigabitethernet
1/0/3
```

```
[DeviceB-rrpp-domain2] ring 2 enable
```

```
[DeviceC-rrpp-domain2] quit
```

# Enable RRPP.

```
[DeviceB] rrpp enable
```

### 3. Configure Device C:

# Create VLANs 1 and 2, map VLAN 1 to MSTI 1 and VLAN 2 to MSTI 2, and activate MST region configuration.

```
<DeviceC> system-view
```

```
[DeviceC] vlan 1 to 2
```

```
[DeviceC] stp region-configuration
```

```
[DeviceC-mst-region] instance 1 vlan 1
```

```
[DeviceC-mst-region] instance 2 vlan 2
```

```
[DeviceC-mst-region] active region-configuration
```

```
[DeviceC-mst-region] quit
```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the two ports as trunk ports, and assign them to VLAN 1 and VLAN 2.

```
[DeviceC] interface gigabitethernet 1/0/1
```

```
[DeviceC-GigabitEthernet1/0/1] undo link-delay
```

```
[DeviceC-GigabitEthernet1/0/1] undo stp enable
```

```
[DeviceC-GigabitEthernet1/0/1] qos trust dot1p
```

```
[DeviceC-GigabitEthernet1/0/1] port link-type trunk
```

```
[DeviceC-GigabitEthernet1/0/1] port trunk permit vlan 1 2
```

```
[DeviceC-GigabitEthernet1/0/1] quit
```

```
[DeviceC] interface gigabitethernet 1/0/2
```

```
[DeviceC-GigabitEthernet1/0/2] undo link-delay
```

```
[DeviceC-GigabitEthernet1/0/2] undo stp enable
```

```
[DeviceC-GigabitEthernet1/0/2] qos trust dot1p
```

```
[DeviceC-GigabitEthernet1/0/2] port link-type trunk
```

```
[DeviceC-GigabitEthernet1/0/2] port trunk permit vlan 1 2
```

```
[DeviceC-GigabitEthernet1/0/2] quit
```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/3, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the port as a trunk port, remove it from VLAN 1, assign it to VLAN 2, and configure VLAN 2 as its default VLAN.

```
[DeviceC] interface gigabitethernet 1/0/3
```

```
[DeviceC-GigabitEthernet1/0/3] undo link-delay
```

```
[DeviceC-GigabitEthernet1/0/3] undo stp enable
```

```

[DeviceC-GigabitEthernet1/0/3] qos trust dot1p
[DeviceC-GigabitEthernet1/0/3] port link-type trunk
[DeviceC-GigabitEthernet1/0/3] undo port trunk permit vlan 1
[DeviceC-GigabitEthernet1/0/3] port trunk permit vlan 2
[DeviceC-GigabitEthernet1/0/3] port trunk pvid vlan 2
[DeviceC-GigabitEthernet1/0/3] quit

Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/4, disable
the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the
port as a trunk port, and assign it to VLAN 1.
[DeviceC] interface gigabitethernet 1/0/4
[DeviceC-GigabitEthernet1/0/4] undo link-delay
[DeviceC-GigabitEthernet1/0/4] undo stp enable
[DeviceC-GigabitEthernet1/0/4] qos trust dot1p
[DeviceC-GigabitEthernet1/0/4] port link-type trunk
[DeviceC-GigabitEthernet1/0/4] port trunk permit vlan 1
[DeviceC-GigabitEthernet1/0/4] quit

Create RRPP domain 1. Configure VLAN 100 as the primary control VLAN of RRPP domain 1,
and configure the VLAN mapped to MSTI 1 as the protected VLAN of RRPP domain 1.
[DeviceC] rrpp domain 1
[DeviceC-rrpp-domain1] control-vlan 100
[DeviceC-rrpp-domain1] protected-vlan reference-instance 1

Configure Device C as the transit node of primary ring 1 in RRPP domain 1, with GigabitEthernet
1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port, and enable ring 1.
[DeviceC-rrpp-domain1] ring 1 node-mode transit primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 0
[DeviceC-rrpp-domain1] ring 1 enable

Configure Device C as the edge node of subring 3 in RRPP domain 1, with GigabitEthernet
1/0/4 as the edge port, and enable subring 3.
[DeviceC-rrpp-domain1] ring 3 node-mode edge edge-port gigabitethernet 1/0/4
[DeviceC-rrpp-domain1] ring 3 enable
[DeviceC-rrpp-domain1] quit

Create RRPP domain 2. Configure VLAN 105 as the primary control VLAN of RRPP domain 2,
and configure the VLAN mapped to MSTI 2 as the protected VLAN of RRPP domain 2.
[DeviceC] rrpp domain 2
[DeviceC-rrpp-domain2] control-vlan 105
[DeviceC-rrpp-domain2] protected-vlan reference-instance 2

Configure Device C as the transit node of primary ring 1 in RRPP domain 2, with GigabitEthernet
1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port, and enable ring 1.
[DeviceC-rrpp-domain2] ring 1 node-mode transit primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 0
[DeviceC-rrpp-domain2] ring 1 enable

Configure Device C as the edge node of subring 2 in RRPP domain 2, with GigabitEthernet
1/0/3 as the edge port, and enable subring 2.
[DeviceC-rrpp-domain2] ring 2 node-mode edge edge-port gigabitethernet 1/0/3
[DeviceC-rrpp-domain2] ring 2 enable
[DeviceC-rrpp-domain2] quit

Enable RRPP.

```

```
[DeviceC] rrpp enable
```

#### 4. Configure Device D:

# Create VLANs 1 and 2, map VLAN 1 to MSTI 1 and VLAN 2 to MSTI 2, and activate MST region configuration.

```
<DeviceD> system-view
[DeviceD] vlan 1 to 2
[DeviceD] stp region-configuration
[DeviceD-mst-region] instance 1 vlan 1
[DeviceD-mst-region] instance 2 vlan 2
[DeviceD-mst-region] active region-configuration
[DeviceD-mst-region] quit
```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the two ports as trunk ports, and assign them to VLAN 1 and VLAN 2.

```
[DeviceD] interface gigabitethernet 1/0/1
[DeviceD-GigabitEthernet1/0/1] undo link-delay
[DeviceD-GigabitEthernet1/0/1] undo stp enable
[DeviceD-GigabitEthernet1/0/1] qos trust dot1p
[DeviceD-GigabitEthernet1/0/1] port link-type trunk
[DeviceD-GigabitEthernet1/0/1] port trunk permit vlan 1 2
[DeviceD-GigabitEthernet1/0/1] quit
[DeviceD] interface gigabitethernet 1/0/2
[DeviceD-GigabitEthernet1/0/2] undo link-delay
[DeviceD-GigabitEthernet1/0/2] undo stp enable
[DeviceD-GigabitEthernet1/0/2] qos trust dot1p
[DeviceD-GigabitEthernet1/0/2] port link-type trunk
[DeviceD-GigabitEthernet1/0/2] port trunk permit vlan 1 2
[DeviceD-GigabitEthernet1/0/2] quit
```

# Create RRPP domain 1. Configure VLAN 100 as the primary control VLAN of RRPP domain 1, and configure the VLAN mapped to MSTI 1 as the protected VLAN of RRPP domain 1.

```
[DeviceD] rrpp domain 1
[DeviceD-rrpp-domain1] control-vlan 100
[DeviceD-rrpp-domain1] protected-vlan reference-instance 1
```

# Configure Device D as the transit node of primary ring 1 in RRPP domain 1, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port, and enable ring 1.

```
[DeviceD-rrpp-domain1] ring 1 node-mode transit primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 0
[DeviceD-rrpp-domain1] ring 1 enable
[DeviceD-rrpp-domain1] quit
```

# Create RRPP domain 2. Configure VLAN 105 as the primary control VLAN of RRPP domain 2, and configure the VLAN mapped to MSTI 2 as the protected VLAN of RRPP domain 2.

```
[DeviceD] rrpp domain 2
[DeviceD-rrpp-domain2] control-vlan 105
[DeviceD-rrpp-domain2] protected-vlan reference-instance 2
```

# Configure Device D as the transit node of primary ring 1 in RRPP domain 2, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port, and enable ring 1.

```
[DeviceD-rrpp-domain2] ring 1 node-mode transit primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 0
[DeviceD-rrpp-domain2] ring 1 enable
[DeviceD-rrpp-domain2] quit
```

**# Enable RRPP.**

```
[DeviceD] rrpp enable
```

## 5. Configure Device E:

**# Create VLAN 2, map VLAN 2 to MSTI 2, and activate MST region configuration.**

```
<DeviceE> system-view
[DeviceE] vlan 2
[DeviceE-vlan2] quit
[DeviceE] stp region-configuration
[DeviceE-mst-region] instance 2 vlan 2
[DeviceE-mst-region] active region-configuration
[DeviceE-mst-region] quit
```

**# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the two ports as trunk ports, remove them from VLAN 1, assign them to VLAN 2, and configure VLAN 2 as their default VLAN.**

```
[DeviceE] interface gigabitethernet 1/0/1
[DeviceE-GigabitEthernet1/0/1] undo link-delay
[DeviceE-GigabitEthernet1/0/1] undo stp enable
[DeviceE-GigabitEthernet1/0/1] qos trust dot1p
[DeviceE-GigabitEthernet1/0/1] port link-type trunk
[DeviceE-GigabitEthernet1/0/1] undo port trunk permit vlan 1
[DeviceE-GigabitEthernet1/0/1] port trunk permit vlan 2
[DeviceE-GigabitEthernet1/0/1] port trunk pvid vlan 2
[DeviceE-GigabitEthernet1/0/1] quit
[DeviceE] interface gigabitethernet 1/0/2
[DeviceE-GigabitEthernet1/0/2] undo link-delay
[DeviceE-GigabitEthernet1/0/2] undo stp enable
[DeviceE-GigabitEthernet1/0/2] qos trust dot1p
[DeviceE-GigabitEthernet1/0/2] port link-type trunk
[DeviceE-GigabitEthernet1/0/2] undo port trunk permit vlan 1
[DeviceE-GigabitEthernet1/0/2] port trunk permit vlan 2
[DeviceE-GigabitEthernet1/0/2] port trunk pvid vlan 2
[DeviceE-GigabitEthernet1/0/2] quit
```

**# Create RRPP domain 2. Configure VLAN 105 as the primary control VLAN, and configure the VLAN mapped to MSTI 2 as the protected VLAN.**

```
[DeviceE] rrpp domain 2
[DeviceE-rrpp-domain2] control-vlan 105
[DeviceE-rrpp-domain2] protected-vlan reference-instance 2
```

**# Configure Device E as the master mode of subring 2 in RRPP domain 2, with GigabitEthernet 1/0/2 as the primary port and GigabitEthernet 1/0/1 as the secondary port, and enable ring 2.**

```
[DeviceE-rrpp-domain2] ring 2 node-mode master primary-port gigabitethernet 1/0/2
secondary-port gigabitethernet 1/0/1 level 1
[DeviceE-rrpp-domain2] ring 2 enable
[DeviceE-rrpp-domain2] quit
```

# Enable RRPP.

```
[DeviceE] rrpp enable
```

## 6. Configure Device F:

# Create VLAN 1, map VLAN 1 to MSTI 1, and activate MST region configuration.

```
<DeviceF> system-view
```

```
[DeviceF] vlan 1
```

```
[DeviceF-vlan1] quit
```

```
[DeviceF] stp region-configuration
```

```
[DeviceF-mst-region] instance 1 vlan 1
```

```
[DeviceF-mst-region] active region-configuration
```

```
[DeviceF-mst-region] quit
```

# Cancel the physical state change suppression interval setting on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, disable the spanning tree feature, and set the trusted packet priority type to 802.1p priority. Configure the two ports as trunk ports, and assign them to VLAN 1.

```
[DeviceF] interface gigabitethernet 1/0/1
```

```
[DeviceF-GigabitEthernet1/0/1] undo link-delay
```

```
[DeviceF-GigabitEthernet1/0/1] undo stp enable
```

```
[DeviceF-GigabitEthernet1/0/1] qos trust dot1p
```

```
[DeviceF-GigabitEthernet1/0/1] port link-type trunk
```

```
[DeviceF-GigabitEthernet1/0/1] port trunk permit vlan 1
```

```
[DeviceF-GigabitEthernet1/0/1] quit
```

```
[DeviceF] interface gigabitethernet 1/0/2
```

```
[DeviceF-GigabitEthernet1/0/2] undo link-delay
```

```
[DeviceF-GigabitEthernet1/0/2] undo stp enable
```

```
[DeviceF-GigabitEthernet1/0/2] qos trust dot1p
```

```
[DeviceF-GigabitEthernet1/0/2] port link-type trunk
```

```
[DeviceF-GigabitEthernet1/0/2] port trunk permit vlan 1
```

```
[DeviceF-GigabitEthernet1/0/2] quit
```

# Create RRPP domain 1. Configure VLAN 100 as the primary control VLAN, and configure the VLAN mapped to MSTI 1 as the protected VLAN.

```
[DeviceF] rrpp domain 1
```

```
[DeviceF-rrpp-domain1] control-vlan 100
```

```
[DeviceF-rrpp-domain1] protected-vlan reference-instance 1
```

# Configure Device F as the master node of subring 3 in RRPP domain 1, with GigabitEthernet 1/0/1 as the primary port and GigabitEthernet 1/0/2 as the secondary port, and enable subring 3.

```
[DeviceF-rrpp-domain1] ring 3 node-mode master primary-port gigabitethernet 1/0/1
secondary-port gigabitethernet 1/0/2 level 1
```

```
[DeviceF-rrpp-domain1] ring 3 enable
```

```
[DeviceF-rrpp-domain1] quit
```

# Enable RRPP.

```
[DeviceF] rrpp enable
```

## 7. RRPP ring group configurations on Device B and Device C after the configurations.

# Create RRPP ring group 1 on Device B. Add subrings 2 and 3 to the RRPP ring group.

```
[DeviceB] rrpp ring-group 1
```

```
[DeviceB-rrpp-ring-group1] domain 2 ring 2
```

```
[DeviceB-rrpp-ring-group1] domain 1 ring 3
```

# Create RRPP ring group 1 on Device C, and add subrings 2 and 3 to the RRPP ring group.

```
[DeviceC] rrpp ring-group 1
[DeviceC-rrpp-ring-group1] domain 2 ring 2
[DeviceC-rrpp-ring-group1] domain 1 ring 3
```

8. Verify the configuration:

Use the **display** command to view RRPP configuration and operational information on each device.

## Troubleshooting

### Symptom

When the link state is normal, the master node cannot receive Hello packets, and the master node unblocks the secondary port.

### Analysis

The reasons may be:

- RRPP is not enabled on some nodes in the RRPP ring.
- The domain ID or primary control VLAN ID is not the same for the nodes in the same RRPP ring.
- Some ports are abnormal.

### Solution

- Use the **display rrpp brief** command to examine whether RRPP is enabled for all nodes. If it is not, use the **rrpp enable** command and the **ring enable** command to enable RRPP and RRPP rings for all nodes.
- Use the **display rrpp brief** command to examine whether the domain ID and primary control VLAN ID are the same for all nodes. If they are not, set the same domain ID and primary control VLAN ID for the nodes.
- Use the **display rrpp verbose** command to examine the link state of each port in each ring.
- Use the **debugging rrpp** command on each node to examine whether a port receives or transmits Hello packets. If it does not, Hello packets are lost.

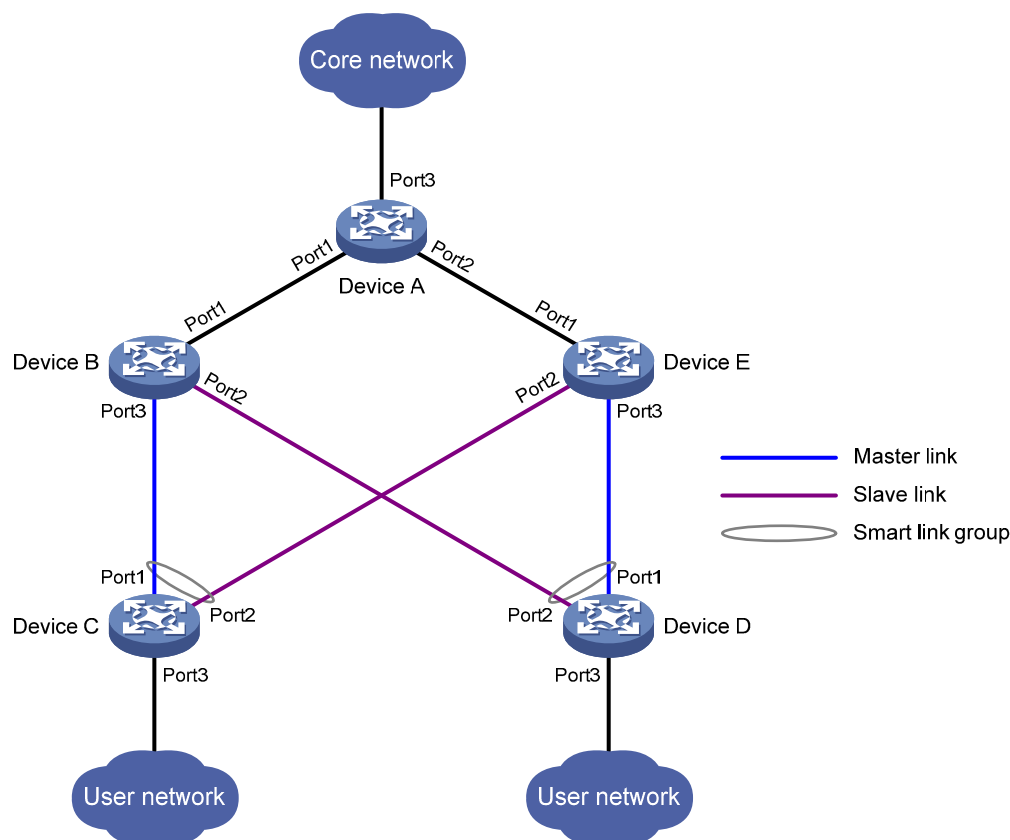
# Configuring Smart Link

## Smart Link overview

### Background

To avoid single-point failures and guarantee network reliability, downstream devices are usually dual-homed to upstream devices, as shown in [Figure 24](#).

**Figure 24 Diagram for a dual uplink network**



To remove network loops on a dual-homed network, you can use a spanning tree protocol or the Rapid Ring Protection Protocol (RRPP). The problem with STP, however, is that STP convergence time is long, which makes it not suitable for users who have high demand on convergence speed. RRPP can meet users' demand on convergence speed, but it involves complicated networking and configurations and is mainly used in ring-shaped networks.

For more information about STP and RRPP, see *Layer 2—LAN Switching Configuration Guide* and "[Configuring RRPP](#)."

Smart Link is a feature developed to address the slow convergence issue with STP. It provides link redundancy as well as fast convergence in a dual uplink network, allowing the backup link to take over quickly when the primary link fails. To sum up, Smart Link has the following features:



- Dedicated to dual uplink networks
- Subsecond convergence
- Easy to configure

## Terminology

### Smart link group

A smart link group consists of only two member ports: the master and the slave ports. At a time, only one port is active for forwarding, and the other port is blocked and in standby state. When link failure occurs on the active port due to port shutdown or presence of unidirectional link for example, the standby port becomes active to take over and the original active port transits to the blocked state.

As shown in [Figure 24](#), Port1 and Port2 of Device C and Port1 and Port2 of Device D each form a smart link group, with Port1 being active and Port2 being standby.

### Master/slave port

Master port and slave port are two port roles in a smart link group. When both ports in a smart link group are up, the master port preferentially transits to the forwarding state, and the slave port stays in standby state. Once the master port fails, the slave port takes over to forward traffic. As shown in [Figure 24](#), you can configure Port1 of Device C and Port1 of Device D as master ports, and Port2 of Device C and Port2 of Device D slave ports.

### Master/slave link

The link that connects the master port in a smart link group is the master link. The link that connects the slave port is the slave link.

### Flush message

Flush messages are used by a smart link group to notify other devices to refresh their MAC address forwarding entries and ARP/ND entries when link switchover occurs in the smart link group. Flush messages are common multicast data packets, and will be dropped by a blocked receiving port.

### Protected VLAN

A smart link group controls the forwarding state of some data VLANs (protected VLANs). Different smart link groups on a port control different protected VLANs. The state of the port in a protected VLAN is determined by the state of the port in the smart link group.

### Transmit control VLAN

The transmit control VLAN is used for transmitting flush messages. When link switchover occurs, the devices (such as Device C and Device D in [Figure 24](#)) broadcast flush messages within the transmit control VLAN.

### Receive control VLAN

The receive control VLAN is used for receiving and processing flush messages. When link switchover occurs, the devices (such as Device A, Device B, and Device E in [Figure 24](#)) receive and process flush messages in the receive control VLAN and refresh their MAC address forwarding entries and ARP/ND entries.

# How Smart Link works

## Link backup mechanism

As shown in [Figure 24](#), the link on Port1 of Device C is the master link, and the link on Port2 of Device C is the slave link. Typically, Port1 is in forwarding state, and Port2 is in standby state. When the master link fails, Port2 takes over to forward traffic and Port1 is blocked and placed in standby state.

---

### NOTE:

When a port switches to the forwarding state, the system outputs log information to notify the user of the port state change.

---

## Topology change mechanism

Because link switchover can outdate the MAC address forwarding entries and ARP/ND entries on all devices, you need a forwarding entry update mechanism to ensure proper transmission. By far, the following two update mechanisms are provided:

- Uplink traffic-triggered MAC address learning, where update is triggered by uplink traffic. This mechanism is applicable to environments with devices not supporting Smart Link, including devices of other vendors'.
- Flush update where a Smart Link-enabled device updates its information by transmitting flush messages over the backup link to its upstream devices. This mechanism requires the upstream devices to be capable of recognizing Smart Link flush messages to update its MAC address forwarding entries and ARP/ND entries.

## Role preemption mechanism

As shown in [Figure 24](#), the link on Port1 of Device C is the master link, and the link on Port2 of Device C is the slave link. Once the master link fails, Port1 is automatically blocked and placed in standby state, and Port2 takes over to forward traffic. When the master link recovers, one of the following occurs:

- If the smart link group is not configured with role preemption, to keep traffic forwarding stable, Port1 that has been blocked due to link failure does not immediately take over to forward traffic. Rather, it stays blocked until the next link switchover.
- If the smart link group is configured with role preemption, Port1 takes over to forward traffic as soon as its link recovers, and Port2 is automatically blocked and placed in standby state.

## Load sharing mechanism

A ring network may carry traffic of multiple VLANs. Smart Link can forward traffic of different VLANs in different smart link groups, implementing load sharing.

To implement load sharing, you can assign a port to multiple smart link groups (each configured with different protected VLANs), making sure that the state of the port is different in these smart link groups. In this way, traffic of different VLANs can be forwarded along different paths.

You can configure protected VLANs for a smart link group by referencing Multiple Spanning Tree Instances (MSTIs).

# Smart Link collaboration mechanisms

## Collaboration between Smart Link and Monitor Link

Smart Link cannot sense by itself when faults occur on the uplink of the upstream devices, or when faults are cleared. To monitor the uplink status of the upstream devices, you can configure the Monitor Link

function to monitor the uplink ports of the upstream devices. Monitor Link adapts the up/down state of downlink ports to the up/down state of uplink ports, triggering Smart Link to perform link switchover on the downstream device.

For more information about Monitor Link, see "[Configuring Monitor Link](#)."

### Collaboration between Smart Link and CC of CFD

Smart Link cannot sense by itself when faults (for example, unidirectional link, misconnected fibers, and packet loss) occur on the intermediate devices or network paths, or when faults are cleared. To check the link status, Smart Link ports must use link detection protocols. When a fault is detected or cleared, the link detection protocols inform Smart Link to switch over the links.

With the collaboration between Smart Link and the Continuity Check (CC) function of Connectivity Fault Detection (CFD) configured, CFD notifies the ports of fault detection events on the basis of detection VLANs and detection ports. A port responds to a continuity check event only when the control VLAN of the smart link group to which it belongs matches the detection VLAN.

For more information about the CC function of CFD, see "[Configuring CFD](#)."

## Smart Link configuration task list

A smart link device is a device that supports Smart Link and is configured with a smart link group and a transmit control VLAN for flush message transmission. Device C and Device D in [Figure 24](#) are two examples of smart link devices.

An associated device is a device that supports Smart Link and receives flush messages sent from the specified control VLAN. Device A, Device B, and Device E in [Figure 24](#) are examples of associated devices.

Complete the following tasks to configure Smart Link:

Task	Remarks	
Configuring a Smart Link device	<a href="#">Configuring protected VLANs for a smart link group</a>	Required
	<a href="#">Configuring member ports for a smart link group</a>	Required
	<a href="#">Configuring role preemption for a smart link group</a>	Optional
	<a href="#">Enabling the sending of flush messages</a>	Optional
	<a href="#">Configuring the collaboration between Smart Link and CC of CFD</a>	Optional
Configuring an associated device	<a href="#">Enabling the receiving of flush messages</a>	Required

## Configuring a Smart Link device

### Configuration prerequisites

- Before configuring a port as a smart link group member, shut down the port to prevent loops. You can bring up the port only after completing the smart link group configuration.

- Disable the spanning tree feature and RRP on the ports that you want to add to the smart link group, and make sure that the ports are not member ports of any aggregation group or service loopback group.

**NOTE:**

A loop may occur on the network during the time when the spanning tree feature is disabled but Smart Link has not yet taken effect on a port.

## Configuring protected VLANs for a smart link group

You can configure protected VLANs for a smart link group by referencing MSTIs. Before configuring the protected VLANs, configure the mappings between MSTIs and the VLANs to be protected. (In PVST mode, the system automatically maps VLANs to MSTIs.) For more information about MSTI and PVST, see *Layer 2—LAN Switching Configuration Guide*.

To configure the protected VLANs for a smart link group:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Enter MST region view.	<b>stp region-configuration</b>	Not required in PVST mode. For more information about the command, see <i>Layer 2—LAN Switching Command Reference</i> .
3. Configure the VLAN-to-instance mapping table.	Approach 1: <b>instance instance-id vlan vlan-list</b> Approach 2: <b>vlan-mapping modulo modulo</b>	Optional. Use either approach. All VLANs in an MST region are mapped to CIST (MSTI 0) by default. Not required in PVST mode. For more information about the commands, see <i>Layer 2—LAN Switching Command Reference</i> .
4. Activate MST region configuration manually.	<b>active region-configuration</b>	Not required in PVST mode. For more information about the command, see <i>Layer 2—LAN Switching Command Reference</i> .
5. Display the currently activated configuration information of the MST region.	<b>display stp region-configuration [   { begin   exclude   include } regular-expression ]</b>	Optional. Available in any view. You can view the VLANs mapped to the MSTIs. For more information about the command, see <i>Layer 2—LAN Switching Command Reference</i> .
6. Return to system view.	<b>quit</b>	Not required in PVST mode.
7. Create a smart link group, and enter smart link group view.	<b>smart-link group group-id</b>	N/A

Step	Command	Remarks
8. Configure protected VLANs for the smart link group.	<b>protected-vlan reference-instance</b> <i>instance-id-list</i>	By default, no protected VLAN is configured for a smart link group.

## Configuring member ports for a smart link group

You can configure member ports for a smart link group either in smart link group view or in interface view. The configurations made in these two views have the same effect.

### In smart link group view

To configure member ports for a smart link group in smart link group view:

Step	Command
1. Enter system view.	<b>system-view</b>
2. Create a smart link group, and enter smart link group view.	<b>smart-link group</b> <i>group-id</i>
3. Configure member ports for a smart link group.	<b>port</b> <i>interface-type interface-number</i> { <b>master</b>   <b>slave</b> }

### In interface view

To configure member ports for a smart link group in interface view:

Step	Command
1. Enter system view.	<b>system-view</b>
2. Enter Layer 2 Ethernet interface view or layer 2 aggregate interface view.	<b>interface</b> <i>interface-type interface-number</i>
3. Configure member ports for a smart link group.	<b>port smart-link group</b> <i>group-id</i> { <b>master</b>   <b>slave</b> }

## Configuring role preemption for a smart link group

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Create a smart link group, and enter smart link group view.	<b>smart-link group</b> <i>group-id</i>	N/A
3. Enable role preemption.	<b>preemption mode role</b>	By default, the device works in the non-preemption mode.
4. Configure the preemption delay.	<b>preemption delay</b> <i>delay-time</i>	Optional. 1 second by default.

#### NOTE:

The preemption delay configuration takes effect only after role preemption is enabled.

## Enabling the sending of flush messages

The control VLAN configured for a smart link group must be different from that configured for any other smart link group.

Make sure the configured control VLAN already exists, and assign the smart link group member ports to the control VLAN.

The control VLAN of a smart link group should also be one of its protected VLANs. Do not remove the control VLAN. Otherwise, flush messages cannot be sent properly.

To enable the sending of flush messages:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Create a smart link group, and enter smart link group view.	<b>smart-link group</b> <i>group-id</i>	N/A
3. Enable flush update in the specified control VLAN.	<b>flush enable</b> [ <b>control-vlan</b> <i>vlan-id</i> ]	Optional. By default, flush update is enabled, and VLAN 1 is the control VLAN.

## Configuring the collaboration between Smart Link and CC of CFD

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Enter Layer 2 Ethernet interface view.	<b>interface</b> <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the collaboration between Smart Link and the CC function of CFD on the port.	<b>port smart-link group</b> <i>group-id</i> <b>track cfd cc</b>	Optional. By default, the collaboration between Smart Link and the CC function of CFD is not configured.

### NOTE:

When configuring the collaboration between Smart Link and the CC function of CFD on a smart link member port, make sure that the control VLAN of the smart link group to which the port belongs matches the detection VLAN of the CC function of CFD.

# Configuring an associated device

## Configuration prerequisites

Disable the spanning tree feature on the associated device's ports that connect to the member ports of the smart link group; otherwise, the ports will discard flush messages when they are not in the forwarding state in case of a topology change.

## Enabling the receiving of flush messages

You do not need to enable all ports on the associated devices to receive flush messages sent from the transmit control VLAN; you only need to enable those on the master and slave links between the smart link device and the destination device.

### Configuration guidelines

- Configure all the control VLANs to receive flush messages.
- If no control VLAN is specified for processing flush messages, the device forwards the received flush messages without processing them.
- Make sure the receive control VLAN is the same as the transmit control VLAN configured on the smart link device. If they are not the same, the associated device will forward the received flush messages directly without any processing.
- Do not remove the control VLANs. Otherwise, flush messages cannot be sent properly.
- Make sure the control VLANs are existing VLANs, and assign the ports capable of receiving flush messages to the control VLANs.

### Configuration procedure

To enable the receiving of flush messages:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view.	<b>interface</b> <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the control VLANs for receiving flush messages.	<b>smart-link flush enable</b> [ <b>control-vlan</b> <i>vlan-id-list</i> ]	By default, no control VLAN exists for receiving flush messages.

## Displaying and maintaining Smart Link

Task	Command	Remarks
Display smart link group information.	<b>display smart-link group</b> { <i>group-id</i>   <b>all</b> } [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]	Available in any view
Display information about the received flush messages.	<b>display smart-link flush</b> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]	Available in any view

Task	Command	Remarks
Clear the statistics about flush messages.	<b>reset smart-link statistics</b>	Available in user view

## Smart Link configuration examples

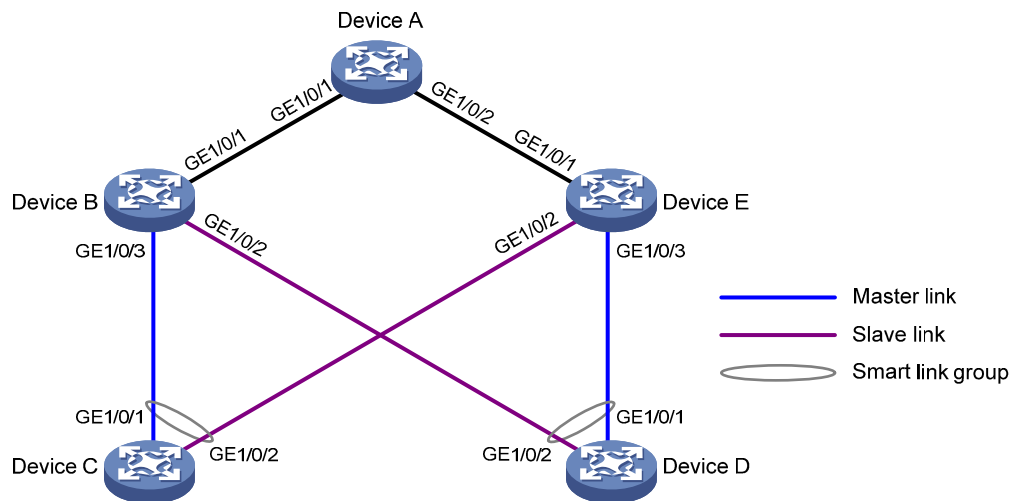
### Single smart link group configuration example

#### Network requirements

As shown in Figure 25, Device C and Device D are smart link devices, and Device A, Device B, and Device E are associated devices. Traffic of VLANs 1 through 30 on Device C and Device D are dually uplinked to Device A.

Configure Smart Link on Device C and Device D for dual uplink backup.

Figure 25 Network diagram



#### Configuration procedure

##### 1. Configure Device C:

# Create VLANs 1 through 30, map these VLANs to MSTI 1, and activate the MST region configuration.

```
<DeviceC> system-view
[DeviceC] vlan 1 to 30
[DeviceC] stp region-configuration
[DeviceC-mst-region] instance 1 vlan 1 to 30
[DeviceC-mst-region] active region-configuration
[DeviceC-mst-region] quit
```

# Shut down GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, disable the spanning tree feature on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 separately, configure them as trunk ports, and assign them to VLANs 1 through 30.

```
[DeviceC] interface gigabitethernet 1/0/1
[DeviceC-GigabitEthernet1/0/1] shutdown
```



```

[DeviceC-GigabitEthernet1/0/1] undo stp enable
[DeviceC-GigabitEthernet1/0/1] port link-type trunk
[DeviceC-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceC-GigabitEthernet1/0/1] quit
[DeviceC] interface gigabitethernet 1/0/2
[DeviceC-GigabitEthernet1/0/2] shutdown
[DeviceC-GigabitEthernet1/0/2] undo stp enable
[DeviceC-GigabitEthernet1/0/2] port link-type trunk
[DeviceC-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceC-GigabitEthernet1/0/2] quit

```

**# Create smart link group 1, and configure all VLANs mapped to MSTI 1 as the protected VLANs.**

```

[DeviceC] smart-link group 1
[DeviceC-smlk-group1] protected-vlan reference-instance 1

```

**# Configure GigabitEthernet 1/0/1 as the master port and GigabitEthernet 1/0/2 as the slave port for smart link group 1.**

```

[DeviceC-smlk-group1] port gigabitethernet1/0/1 master
[DeviceC-smlk-group1] port gigabitethernet1/0/2 slave

```

**# Enable flush message sending in smart link group 1, and configure VLAN 10 as the transmit control VLAN.**

```

[DeviceC-smlk-group1] flush enable control-vlan 10
[DeviceC-smlk-group1] quit

```

**# Bring up GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 again.**

```

[DeviceC] interface gigabitethernet 1/0/1
[DeviceC-GigabitEthernet1/0/1] undo shutdown
[DeviceC-GigabitEthernet1/0/1] quit
[DeviceC] interface gigabitethernet 1/0/2
[DeviceC-GigabitEthernet1/0/2] undo shutdown
[DeviceC-GigabitEthernet1/0/2] quit

```

## 2. Configure Device D:

**# Create VLANs 1 through 30, map these VLANs to MSTI 1, and activate the MST region configuration.**

```

<DeviceD> system-view
[DeviceD] vlan 1 to 30
[DeviceD] stp region-configuration
[DeviceD-mst-region] instance 1 vlan 1 to 30
[DeviceD-mst-region] active region-configuration
[DeviceD-mst-region] quit

```

**# Shut down GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, disable the spanning tree feature on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 separately, configure them as trunk ports, and assign them to VLANs 1 through 30.**

```

[DeviceD] interface gigabitethernet 1/0/1
[DeviceD-GigabitEthernet1/0/1] shutdown
[DeviceD-GigabitEthernet1/0/1] undo stp enable
[DeviceD-GigabitEthernet1/0/1] port link-type trunk
[DeviceD-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceD-GigabitEthernet1/0/1] quit
[DeviceD] interface gigabitethernet 1/0/2

```

```

[DeviceD-GigabitEthernet1/0/2] shutdown
[DeviceD-GigabitEthernet1/0/2] undo stp enable
[DeviceD-GigabitEthernet1/0/2] port link-type trunk
[DeviceD-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceD-GigabitEthernet1/0/2] quit
Create smart link group 1, and configure all VLANs mapped to MSTI 1 as the protected VLANs.
[DeviceD] smart-link group 1
[DeviceD-smlk-group1] protected-vlan reference-instance 1
Configure GigabitEthernet 1/0/1 as the master port and GigabitEthernet 1/0/2 as the slave
port for smart link group 1.
[DeviceD-smlk-group1] port gigabitethernet1/0/1 master
[DeviceD-smlk-group1] port gigabitethernet1/0/2 slave
Enable flush message sending in smart link group 1, and configure VLAN 20 as the transmit
control VLAN.
[DeviceD-smlk-group1] flush enable control-vlan 20
[DeviceD-smlk-group1] quit
Bring up GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 again.
[DeviceD] interface gigabitethernet 1/0/1
[DeviceD-GigabitEthernet1/0/1] undo shutdown
[DeviceD-GigabitEthernet1/0/1] quit
[DeviceD] interface gigabitethernet 1/0/2
[DeviceD-GigabitEthernet1/0/2] undo shutdown
[DeviceD-GigabitEthernet1/0/2] quit

```

### 3. Configure Device B:

# Create VLANs 1 through 30.

```

<DeviceB> system-view
[DeviceB] vlan 1 to 30

```

# Configure GigabitEthernet 1/0/1 as a trunk port, and assign it to VLANs 1 through 30. Enable flush message receiving on it, and configure VLAN 10 and VLAN 20 as the receive control VLANs..

```

[DeviceB] interface gigabitethernet 1/0/1
[DeviceB-GigabitEthernet1/0/1] port link-type trunk
[DeviceB-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceB-GigabitEthernet1/0/1] smart-link flush enable control-vlan 10 20
[DeviceB-GigabitEthernet1/0/1] quit

```

# Configure GigabitEthernet 1/0/2 as a trunk port, and assign it to VLANs 1 through 30. Disable the spanning tree feature and enable flush message receiving on it, and configure VLAN 20 as the receive control VLAN.

```

[DeviceB] interface gigabitethernet 1/0/2
[DeviceB-GigabitEthernet1/0/2] port link-type trunk
[DeviceB-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceB-GigabitEthernet1/0/2] undo stp enable
[DeviceB-GigabitEthernet1/0/2] smart-link flush enable control-vlan 20
[DeviceB-GigabitEthernet1/0/2] quit

```

# Configure GigabitEthernet 1/0/3 as a trunk port, and assign it to VLANs 1 through 30. Disable the spanning tree feature and enable flush message receiving on it, and configure VLAN 10 as the receive control VLAN.

```
[DeviceB] interface gigabitethernet 1/0/3
[DeviceB-GigabitEthernet1/0/3] port link-type trunk
[DeviceB-GigabitEthernet1/0/3] port trunk permit vlan 1 to 30
[DeviceB-GigabitEthernet1/0/3] undo stp enable
[DeviceB-GigabitEthernet1/0/3] smart-link flush enable control-vlan 10
[DeviceB-GigabitEthernet1/0/3] quit
```

#### 4. Configure Device E:

# Create VLANs 1 through 30.

```
<DeviceE> system-view
[DeviceE] vlan 1 to 30
```

# Configure GigabitEthernet 1/0/1 as a trunk port, and assign it to VLANs 1 through 30. Enable flush message receiving on it, and configure VLAN 10 and VLAN 20 as the receive control VLANs.

```
[DeviceE] interface gigabitethernet 1/0/1
[DeviceE-GigabitEthernet1/0/1] port link-type trunk
[DeviceE-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceE-GigabitEthernet1/0/1] smart-link flush enable control-vlan 10 20
[DeviceE-GigabitEthernet1/0/1] quit
```

# Configure GigabitEthernet 1/0/2 as a trunk port, and assign it to VLANs 1 through 30. Disable the spanning tree feature and enable flush message receiving on it, and configure VLAN 10 as the receive control VLAN.

```
[DeviceE] interface gigabitethernet 1/0/2
[DeviceE-GigabitEthernet1/0/2] port link-type trunk
[DeviceE-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceE-GigabitEthernet1/0/2] undo stp enable
[DeviceE-GigabitEthernet1/0/2] smart-link flush enable control-vlan 10
[DeviceE-GigabitEthernet1/0/2] quit
```

# Configure GigabitEthernet 1/0/3 as a trunk port, and assign it to VLANs 1 through 30. Disable the spanning tree feature and enable flush message receiving on it, and configure VLAN 20 as the receive control VLAN.

```
[DeviceE] interface gigabitethernet 1/0/3
[DeviceE-GigabitEthernet1/0/3] port link-type trunk
[DeviceE-GigabitEthernet1/0/3] port trunk permit vlan 1 to 30
[DeviceE-GigabitEthernet1/0/3] undo stp enable
[DeviceE-GigabitEthernet1/0/3] smart-link flush enable control-vlan 20
[DeviceE-GigabitEthernet1/0/3] quit
```

#### 5. Configure Device A:

# Create VLANs 1 through 30.

```
<DeviceA> system-view
[DeviceA] vlan 1 to 30
```

# Configure GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 as trunk ports, and assign them to VLANs 1 through 30. Enable flush message receiving on them, and configure VLAN 10 and VLAN 20 as the receive control VLANs.

```
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] port link-type trunk
[DeviceA-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceA-GigabitEthernet1/0/1] smart-link flush enable control-vlan 10 20
[DeviceA-GigabitEthernet1/0/1] quit
```

```
[DeviceA] interface gigabitethernet 1/0/2
[DeviceA-GigabitEthernet1/0/2] port link-type trunk
[DeviceA-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceA-GigabitEthernet1/0/2] smart-link flush enable control-vlan 10 20
[DeviceA-GigabitEthernet1/0/2] quit
```

## 6. Verify the configuration:

You can use the **display smart-link group** command to display the smart link group configuration on a device.

# Display the smart link group configuration on Device C.

```
[DeviceC] display smart-link group 1
Smart link group 1 information:
Device ID: 000f-e23d-5af0
Preemption mode: NONE
Preemption delay: 1(s)
Control VLAN: 10
Protected VLAN: Reference Instance 1
```

Member	Role	State	Flush-count	Last-flush-time
GigabitEthernet1/0/1	MASTER	ACTVIE	5	16:37:20 2010/02/21
GigabitEthernet1/0/2	SLAVE	STANDBY	1	17:45:20 2010/02/21

You can use the **display smart-link flush** command to display the flush messages received on a device.

# Display the flush messages received on Device B.

```
[DeviceB] display smart-link flush
Received flush packets : 5
Receiving interface of the last flush packet : GigabitEthernet1/0/3
Receiving time of the last flush packet : 16:25:21 2009/02/21
Device ID of the last flush packet : 000f-e23d-5af0
Control VLAN of the last flush packet : 10
```

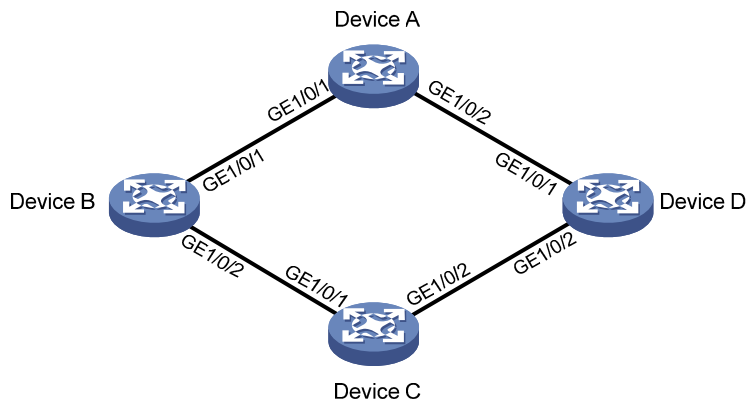
## Multiple smart link groups load sharing configuration example

### Network requirements

As shown in [Figure 26](#), Device C is a smart link device, and Device A, Device B, and Device D are associated devices. Traffic of VLANs 1 through 200 on Device C are dually uplinked to Device A by Device B and Device D.

Implement dual uplink backup and load sharing on Device C: traffic of VLANs 1 through 100 is uplinked to Device A by Device B; traffic of VLANs 101 through 200 is uplinked to Device A by Device D.

Figure 26 Network diagram



## Configuration procedure

### 1. Configure Device C:

# Create VLAN 1 through VLAN 200. Map VLANs 1 through 100 to MSTI 1. Map VLANs 101 through 200 to MSTI 2, and activate MST region configuration.

```
<DeviceC> system-view
[DeviceC] vlan 1 to 200
[DeviceC] stp region-configuration
[DeviceC-mst-region] instance 1 vlan 1 to 100
[DeviceC-mst-region] instance 2 vlan 101 to 200
[DeviceC-mst-region] active region-configuration
[DeviceC-mst-region] quit
```

# Shut down GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, disable the spanning tree feature on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 separately, configure the ports as trunk ports, and assign them to VLAN 1 through VLAN 200.

```
[DeviceC] interface gigabitethernet 1/0/1
[DeviceC-GigabitEthernet1/0/1] shutdown
[DeviceC-GigabitEthernet1/0/1] undo stp enable
[DeviceC-GigabitEthernet1/0/1] port link-type trunk
[DeviceC-GigabitEthernet1/0/1] port trunk permit vlan 1 to 200
[DeviceC-GigabitEthernet1/0/1] quit
[DeviceC] interface gigabitethernet 1/0/2
[DeviceC-GigabitEthernet1/0/2] shutdown
[DeviceC-GigabitEthernet1/0/2] undo stp enable
[DeviceC-GigabitEthernet1/0/2] port link-type trunk
[DeviceC-GigabitEthernet1/0/2] port trunk permit vlan 1 to 200
[DeviceC-GigabitEthernet1/0/2] quit
```

# Create smart link group 1, and configure all VLANs mapped to MSTI 1 as the protected VLANs for smart link group 1.

```
[DeviceC] smart-link group 1
[DeviceC-smlk-group1] protected-vlan reference-instance 1
```

# Configure GigabitEthernet 1/0/1 as the master port and GigabitEthernet 1/0/2 as the slave port for smart link group 1.

```
[DeviceC-smlk-group1] port gigabitethernet1/0/1 master
[DeviceC-smlk-group1] port gigabitethernet1/0/2 slave
```

# Enable role preemption in smart link group 1, enable flush message sending, and configure VLAN 10 as the transmit control VLAN.

```
[DeviceC-smlk-group1] preempt mode role
[DeviceC-smlk-group-1] flush enable control-vlan 10
[DeviceC-smlk-group-1] quit
```

# Create smart link group 2, and configure all VLANs mapped to MSTI 2 as the protected VLANs for smart link group 2.

```
[DeviceC] smart-link group 2
[DeviceC-smlk-group2] protected-vlan reference-instance 2
```

# Configure GigabitEthernet 1/0/1 as the slave port and GigabitEthernet 1/0/2 as the master port for smart link group 2.

```
[DeviceC-smlk-group2] port gigabitethernet1/0/2 master
[DeviceC-smlk-group2] port gigabitethernet1/0/1 slave
```

# Enable role preemption in smart link group 2, enable flush message sending, and configure VLAN 110 as the transmit control VLAN.

```
[DeviceC-smlk-group2] preempt mode role
[DeviceC-smlk-group2] flush enable control-vlan 110
[DeviceC-smlk-group2] quit
```

# Bring up GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 again.

```
[DeviceC] interface gigabitethernet 1/0/1
[DeviceC-GigabitEthernet1/0/1] undo shutdown
[DeviceC-GigabitEthernet1/0/1] quit
[DeviceC] interface gigabitethernet 1/0/2
[DeviceC-GigabitEthernet1/0/2] undo shutdown
[DeviceC-GigabitEthernet1/0/2] quit
```

## 2. Configure Device B:

# Create VLAN 1 through VLAN 200.

```
<DeviceB> system-view
[DeviceB] vlan 1 to 200
```

# Configure GigabitEthernet 1/0/1 as a trunk port and assign it to VLANs 1 through 200. Enable flush message receiving and configure VLAN 10 and VLAN 110 as the receive control VLANs on GigabitEthernet 1/0/1.

```
[DeviceB] interface gigabitethernet 1/0/1
[DeviceB-GigabitEthernet1/0/1] port link-type trunk
[DeviceB-GigabitEthernet1/0/1] port trunk permit vlan 1 to 200
[DeviceB-GigabitEthernet1/0/1] smart-link flush enable control-vlan 10 110
[DeviceB-GigabitEthernet1/0/1] quit
```

# Configure GigabitEthernet 1/0/2 as a trunk port and assign it to VLANs 1 through 200. Disable the spanning tree feature and enable flush message receiving on it, and configure VLAN 10 and VLAN 110 as the receive control VLANs.

```
[DeviceB] interface gigabitethernet 1/0/2
[DeviceB-GigabitEthernet1/0/2] port link-type trunk
[DeviceB-GigabitEthernet1/0/2] port trunk permit vlan 1 to 200
[DeviceB-GigabitEthernet1/0/2] undo stp enable
[DeviceB-GigabitEthernet1/0/2] smart-link flush enable control-vlan 10 110
[DeviceB-GigabitEthernet1/0/2] quit
```

## 3. Configure Device D:

# Create VLAN 1 through VLAN 200.

```
<DeviceD> system-view
[DeviceD] vlan 1 to 200
```

# Configure GigabitEthernet 1/0/1 as a trunk port and assign it to VLANs 1 through 200. Enable flush message receiving and configure VLAN 10 and VLAN 110 as the receive control VLANs on GigabitEthernet 1/0/1.

```
[DeviceD] interface gigabitethernet 1/0/1
[DeviceD-GigabitEthernet1/0/1] port link-type trunk
[DeviceD-GigabitEthernet1/0/1] port trunk permit vlan 1 to 200
[DeviceD-GigabitEthernet1/0/1] smart-link flush enable control-vlan 10 110
[DeviceD-GigabitEthernet1/0/1] quit
```

# Configure GigabitEthernet 1/0/2 as a trunk port and assign it to VLANs 1 through 200. Disable the spanning tree feature and enable flush message receiving on it, and configure VLAN 10 and VLAN 110 as the receive control VLANs.

```
[DeviceD] interface gigabitethernet 1/0/2
[DeviceD-GigabitEthernet1/0/2] port link-type trunk
[DeviceD-GigabitEthernet1/0/2] port trunk permit vlan 1 to 200
[DeviceD-GigabitEthernet1/0/2] undo stp enable
[DeviceD-GigabitEthernet1/0/2] smart-link flush enable control-vlan 10 110
[DeviceD-GigabitEthernet1/0/2] quit
```

#### 4. Configure Device A:

# Create VLAN 1 through VLAN 200.

```
<DeviceA> system-view
[DeviceA] vlan 1 to 200
```

# Configure GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 as trunk ports and assign them to VLANs 1 through 200. Enable flush message receiving and configure VLAN 10 and VLAN 110 as the receive control VLANs on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2.

```
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] port link-type trunk
[DeviceA-GigabitEthernet1/0/1] port trunk permit vlan 1 to 200
[DeviceA-GigabitEthernet1/0/1] smart-link flush enable control-vlan 10 110
[DeviceA-GigabitEthernet1/0/1] quit
[DeviceA] interface gigabitethernet 1/0/2
[DeviceA-GigabitEthernet1/0/2] port link-type trunk
[DeviceA-GigabitEthernet1/0/2] port trunk permit vlan 1 to 200
[DeviceA-GigabitEthernet1/0/2] smart-link flush enable control-vlan 10 110
[DeviceA-GigabitEthernet1/0/2] quit
```

#### 5. Verify the configuration:

You can use the **display smart-link group** command to display the smart link group configuration on a device.

# Display the smart link group configuration on Device C.

```
[DeviceC] display smart-link group all
Smart link group 1 information:
Device ID: 000f-e23d-5af0
Preemption delay: 1(s)
Preemption mode: ROLE
Control VLAN: 10
```

```

Protected VLAN: Reference Instance 1
Member Role State Flush-count Last-flush-time

GigabitEthernet1/0/1 MASTER ACTVIE 5 16:37:20 2010/02/21

GigabitEthernet1/0/2 SLAVE STANDBY 1 17:45:20 2010/02/21

```

Smart link group 2 information:

Device ID: 000f-e23d-5af0

Preemption mode: ROLE

Preemption delay: 1(s)

Control VLAN: 110

```

Protected VLAN: Reference Instance 2
Member Role State Flush-count Last-flush-time

GigabitEthernet1/0/2 MASTER ACTVIE 5 16:37:20 2010/02/21

GigabitEthernet1/0/1 SLAVE STANDBY 1 17:45:20 2010/02/21

```

You can use the **display smart-link flush** command to display the flush messages received on a device.

# Display the flush messages received on Device B.

```

[DeviceB] display smart-link flush
Received flush packets : 5
Receiving interface of the last flush packet : GigabitEthernet1/0/2
Receiving time of the last flush packet : 16:25:21 2010/02/21
Device ID of the last flush packet : 000f-e23d-5af0
Control VLAN of the last flush packet : 10

```

## Smart Link and CFD collaboration configuration example

### Network requirements

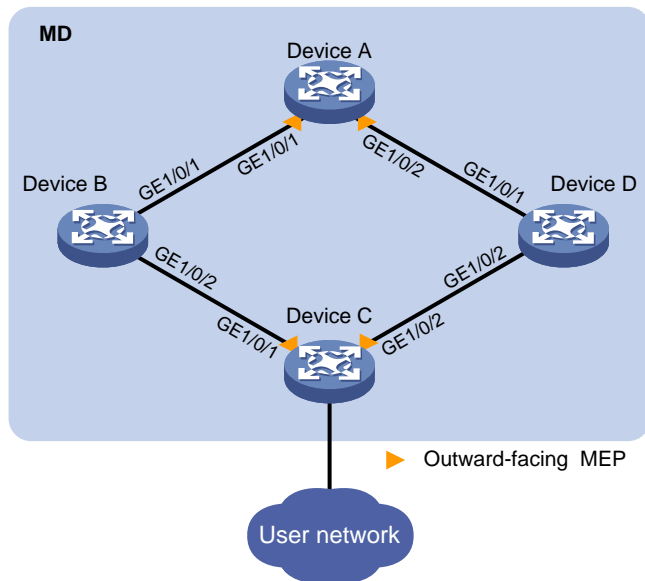
As shown in [Figure 27](#), Device A, Device B, Device C, and Device D form a maintenance domain (MD) of level 5. Device C is a smart link device, and Device A, Device B, and Device D are associated devices. Traffic of VLANs 1 through 200 on Device C is dually uplinked to Device A by Device B and Device D.

Configure the CFD CC function for Smart Link, so that; Traffic of VLANs 1 through 100 is uplinked to Device A by Device C through GigabitEthernet 1/0/1 (master port of smart link group 1). Traffic of VLANs 101 through 200 is uplinked to Device A by Device C through GigabitEthernet 1/0/2 (master port of smart link group 2). When the link between Device C and Device A fails, traffic is rapidly switched to the slave port of each smart link group, and switched back to the master ports after the fault is cleared.

For more information about CFD, see "[Configuring CFD](#)."



Figure 27 Network diagram



## Configuration procedure

### 1. Configure Device A:

# Create VLAN 1 through VLAN 200.

```
<DeviceA> system-view
[DeviceA] vlan 1 to 200
```

# Configure GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 as trunk ports and assign them to VLANs 1 through 200. Enable flush message receiving and configure VLAN 10 and VLAN 110 as the receive control VLANs on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2.

```
[DeviceA] interface gigabitethernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] port link-type trunk
[DeviceA-GigabitEthernet1/0/1] port trunk permit vlan 1 to 200
[DeviceA-GigabitEthernet1/0/1] smart-link flush enable control-vlan 10 110
[DeviceA-GigabitEthernet1/0/1] quit
[DeviceA] interface gigabitethernet 1/0/2
[DeviceA-GigabitEthernet1/0/2] port link-type trunk
[DeviceA-GigabitEthernet1/0/2] port trunk permit vlan 1 to 200
[DeviceA-GigabitEthernet1/0/2] smart-link flush enable control-vlan 10 110
[DeviceA-GigabitEthernet1/0/2] quit
```

# Enable CFD and create an MD of level 5.

```
[DeviceA] cfd enable
[DeviceA] cfd md MD level 5
```

# Create MA **MA\_A** for the MD and configure the MA to serve VLAN 10, and create service instance 1 for the MD and MA.

```
[DeviceA] cfd ma MA_A md MD vlan 10
[DeviceA] cfd service-instance 1 md MD ma MA_A
```

# Create a MEP list in service instance 1, create and enable outward-facing MEP 1002, and enable CCM sending in service instance 1 on GigabitEthernet 1/0/1.

```
[DeviceA] cfd meplist 1001 1002 service-instance 1
[DeviceA] interface gigabitethernet 1/0/1
```

```
[DeviceA-GigabitEthernet1/0/1] cfd mep 1002 service-instance 1 outbound
[DeviceA-GigabitEthernet1/0/1] cfd mep service-instance 1 mep 1002 enable
[DeviceA-GigabitEthernet1/0/1] cfd cc service-instance 1 mep 1002 enable
[DeviceA-GigabitEthernet1/0/1] quit
```

# Create MA **MA\_B** for the MD and configure the MA to serve VLAN 110, and create service instance 2 for the MD and MA.

```
[DeviceA] cfd ma MA_B md MD vlan 110
[DeviceA] cfd service-instance 2 md MD ma MA_B
```

# Create a MEP list in service instance 2, create and enable outward-facing MEP 1002, and enable CCM sending in service instance 2 on GigabitEthernet 1/0/2.

```
[DeviceA] cfd meplist 2001 2002 service-instance 2
[DeviceA] interface gigabitethernet 1/0/2
[DeviceA-GigabitEthernet1/0/2] cfd mep 2002 service-instance 2 outbound
[DeviceA-GigabitEthernet1/0/2] cfd mep service-instance 2 mep 2002 enable
[DeviceA-GigabitEthernet1/0/2] cfd cc service-instance 2 mep 2002 enable
[DeviceA-GigabitEthernet1/0/2] quit
```

## 2. Configure Device B:

# Create VLAN 1 through VLAN 200.

```
<DeviceB> system-view
[DeviceB] vlan 1 to 200
```

# Configure GigabitEthernet 1/0/1 as a trunk port and assign it to VLANs 1 through 200. Enable flush message receiving and configure VLAN 10 and VLAN 110 as the receive control VLANs on GigabitEthernet 1/0/1.

```
[DeviceB] interface gigabitethernet 1/0/1
[DeviceB-GigabitEthernet1/0/1] port link-type trunk
[DeviceB-GigabitEthernet1/0/1] port trunk permit vlan 1 to 200
[DeviceB-GigabitEthernet1/0/1] smart-link flush enable control-vlan 10 110
[DeviceB-GigabitEthernet1/0/1] quit
```

# Configure GigabitEthernet 1/0/2 as a trunk port and assign it to VLANs 1 through 200. Disable the spanning tree feature and enable flush message receiving on it, and configure VLAN 10 and VLAN 110 as the receive control VLANs.

```
[DeviceB] interface gigabitethernet 1/0/2
[DeviceB-GigabitEthernet1/0/2] port link-type trunk
[DeviceB-GigabitEthernet1/0/2] port trunk permit vlan 1 to 200
[DeviceB-GigabitEthernet1/0/2] undo stp enable
[DeviceB-GigabitEthernet1/0/2] smart-link flush enable control-vlan 10 110
[DeviceB-GigabitEthernet1/0/2] quit
```

## 3. Configure Device C:

# Create VLAN 1 through VLAN 200, map VLANs 1 through 100 to MSTI 1, and VLANs 101 through 200 to MSTI 2, and activate MST region configuration.

```
<DeviceC> system-view
[DeviceC] vlan 1 to 200
[DeviceC] stp region-configuration
[DeviceC-mst-region] instance 1 vlan 1 to 100
[DeviceC-mst-region] instance 2 vlan 101 to 200
[DeviceC-mst-region] active region-configuration
[DeviceC-mst-region] quit
```

# Shut down GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2, disable the spanning tree feature on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 separately, configure the ports as trunk ports, and assign them to VLAN 1 through VLAN 200.

```
[DeviceC] interface gigabitethernet 1/0/1
[DeviceC-GigabitEthernet1/0/1] shutdown
[DeviceC-GigabitEthernet1/0/1] undo stp enable
[DeviceC-GigabitEthernet1/0/1] port link-type trunk
[DeviceC-GigabitEthernet1/0/1] port trunk permit vlan 1 to 200
[DeviceC-GigabitEthernet1/0/1] quit
[DeviceC] interface gigabitethernet 1/0/2
[DeviceC-GigabitEthernet1/0/2] shutdown
[DeviceC-GigabitEthernet1/0/2] undo stp enable
[DeviceC-GigabitEthernet1/0/2] port link-type trunk
[DeviceC-GigabitEthernet1/0/2] port trunk permit vlan 1 to 200
[DeviceC-GigabitEthernet1/0/2] quit
```

# Create smart link group 1, and configure all VLANs mapped to MSTI 1 as the protected VLANs for smart link group 1.

```
[DeviceC] smart-link group 1
[DeviceC-smlk-group1] protected-vlan reference-instance 1
```

# Configure GigabitEthernet 1/0/1 as the master port and GigabitEthernet 1/0/2 as the slave port for smart link group 1.

```
[DeviceC-smlk-group1] port gigabitethernet1/0/1 master
[DeviceC-smlk-group1] port gigabitethernet1/0/2 slave
```

# Enable role preemption in smart link group 1, enable flush message sending, and configure VLAN 10 as the transmit control VLAN.

```
[DeviceC-smlk-group1] preemption mode role
[DeviceC-smlk-group1] flush enable control-vlan 10
[DeviceC-smlk-group1] quit
```

# Create smart link group 2, and configure all VLANs mapped to MSTI 2 as the protected VLANs for smart link group 2.

```
[DeviceC] smart-link group 2
[DeviceC-smlk-group2] protected-vlan reference-instance 2
```

# Configure GigabitEthernet 1/0/1 as the slave port and GigabitEthernet 1/0/2 as the master port for smart link group 2.

```
[DeviceC-smlk-group2] port gigabitethernet1/0/2 master
[DeviceC-smlk-group2] port gigabitethernet1/0/1 slave
```

# Enable role preemption in smart link group 2, enable flush message sending, and configure VLAN 110 as the transmit control VLAN.

```
[DeviceC-smlk-group2] preemption mode role
[DeviceC-smlk-group2] flush enable control-vlan 110
[DeviceC-smlk-group2] quit
```

# Enable CFD and create an MD of level 5.

```
[DeviceC] cfd enable
[DeviceC] cfd md MD level 5
```

# Create MA **MA\_A** for the MD and configure the MA to serve VLAN 10, and create service instance 1 for the MD and MA.

```
[DeviceC] cfd ma MA_A md MD vlan 10
```

```
[DeviceC] cfd service-instance 1 md MD ma MA_A
Create a MEP list in service instance 1, create and enable outward-facing MEP 1001, and
enable CCM sending in service instance 1 on GigabitEthernet 1/0/1.
```

```
[DeviceC] cfd meplist 1001 1002 service-instance 1
[DeviceC] interface gigabitethernet 1/0/1
[DeviceC-GigabitEthernet1/0/1] cfd mep 1001 service-instance 1 outbound
[DeviceC-GigabitEthernet1/0/1] cfd mep service-instance 1 mep 1001 enable
[DeviceC-GigabitEthernet1/0/1] cfd cc service-instance 1 mep 1001 enable
[DeviceC-GigabitEthernet1/0/1] quit
```

# Create MA **MA\_B** for the MD and configure the MA to serve VLAN 110, and create service instance 2 for the MD and MA.

```
[DeviceC] cfd ma MA_B md MD vlan 110
[DeviceC] cfd service-instance 2 md MD ma MA_B
```

# Create a MEP list in service instance 2, create and enable outward-facing MEP 2001, and enable CCM sending in service instance 2 on GigabitEthernet 1/0/2.

```
[DeviceC] cfd meplist 2001 2002 service-instance 2
[DeviceC] interface gigabitethernet 1/0/2
[DeviceC-GigabitEthernet1/0/2] cfd mep 2001 service-instance 2 outbound
[DeviceC-GigabitEthernet1/0/2] cfd mep service-instance 2 mep 2001 enable
[DeviceC-GigabitEthernet1/0/2] cfd cc service-instance 2 mep 2001 enable
[DeviceC-GigabitEthernet1/0/2] quit
```

# Configure the collaboration between the master port GigabitEthernet 1/0/1 of smart link group 1 and the CC function of CFD, and bring up the port.

```
[DeviceC] interface gigabitethernet 1/0/1
[DeviceC-GigabitEthernet1/0/1] port smart-link group 1 track cfd cc
[DeviceC-GigabitEthernet1/0/1] undo shutdown
[DeviceC-GigabitEthernet1/0/1] quit
```

# Configure the collaboration between the master port GigabitEthernet 1/0/2 of smart link group 2 and the CC function of CFD, and bring up the port.

```
[DeviceC] interface gigabitethernet 1/0/2
[DeviceC-GigabitEthernet1/0/1] port smart-link group 2 track cfd cc
[DeviceC-GigabitEthernet1/0/2] undo shutdown
[DeviceC-GigabitEthernet1/0/2] quit
```

#### 4. Configure Device D:

# Create VLAN 1 through VLAN 200.

```
<DeviceD> system-view
[DeviceD] vlan 1 to 200
```

# Configure GigabitEthernet 1/0/1 as a trunk port and assign it to VLANs 1 through 200. Enable flush message receiving and configure VLAN 10 and VLAN 110 as the receive control VLANs on GigabitEthernet 1/0/1.

```
[DeviceD] interface gigabitethernet 1/0/1
[DeviceD-GigabitEthernet1/0/1] port link-type trunk
[DeviceD-GigabitEthernet1/0/1] port trunk permit vlan 1 to 200
[DeviceD-GigabitEthernet1/0/1] smart-link flush enable control-vlan 10 110
[DeviceD-GigabitEthernet1/0/1] quit
```

# Configure GigabitEthernet 1/0/2 as a trunk port and assign it to VLANs 1 through 200. Disable the spanning tree feature and enable flush message receiving on it, and configure VLAN 10 and VLAN 110 as the receive control VLANs.

```
[DeviceD] interface gigabitethernet 1/0/2
[DeviceD-GigabitEthernet1/0/2] port link-type trunk
[DeviceD-GigabitEthernet1/0/2] port trunk permit vlan 1 to 200
[DeviceD-GigabitEthernet1/0/2] undo stp enable
[DeviceD-GigabitEthernet1/0/2] smart-link flush enable control-vlan 10 110
[DeviceD-GigabitEthernet1/0/2] quit
```

5. Verify the configuration:

Suppose the optical fiber between Device A and Device B fails. You can use the **display smart-link group** command to display the smart link group configuration on a device.

# Display the smart link group configuration on Device C.

```
[DeviceC] display smart-link group all
Smart link group 1 information:
Device ID: 000f-e23d-5af0
Preemption mode: ROLE
Preemption delay: 1(s)
Control VLAN: 10
Protected VLAN: Reference Instance 1
Member Role State Flush-count Last-flush-time

GigabitEthernet1/0/1 MASTER DOWN 5 16:37:20 2010/02/21
GigabitEthernet1/0/2 SLAVE ACTVIE 3 17:45:20 2010/02/21

Smart link group 2 information:
Device ID: 000f-e23d-5af0
Preemption mode: ROLE
Preemption delay: 1(s)
Control VLAN: 110
Protected VLAN: Reference Instance 2
Member Role State Flush-count Last-flush-time

GigabitEthernet1/0/2 MASTER ACTVIE 5 16:37:20 2010/02/21
GigabitEthernet1/0/1 SLAVE STANDBY 1 17:45:20 2010/02/21
```

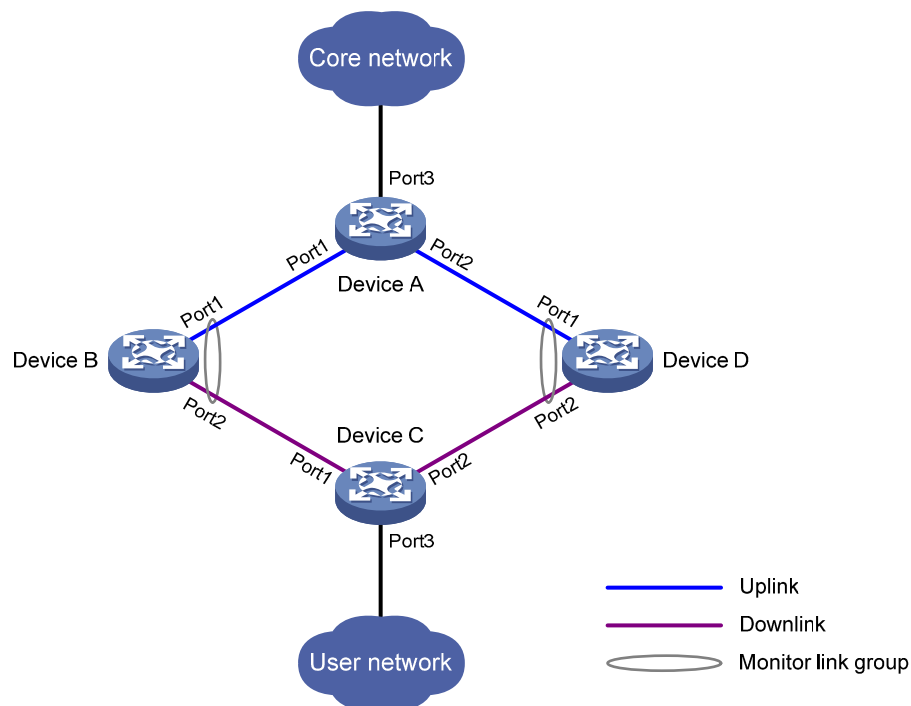
The output shows that master port GigabitEthernet 1/0/1 of smart link group 1 fails, and slave port GigabitEthernet 1/0/2 is in forwarding state.

# Configuring Monitor Link

## Monitor Link overview

Monitor Link is a port collaboration function. Monitor Link usually works together with Layer 2 topology protocols. The idea is to monitor the states of uplink ports and adapt the up/down state of downlink ports to the up/down state of uplink ports, triggering link switchover on the downstream device in time, as shown in [Figure 28](#).

**Figure 28 Monitor Link application scenario**



## Terminology

### Monitor link group

A monitor link group is a set of uplink and downlink ports. A port can belong to only one monitor link group. As shown in [Figure 28](#), ports Port1 and Port2 of Device B and those of Device D each form a monitor link group. Port1 on both devices are uplink ports, and Port2 on both devices are downlink ports.

### Uplink/Downlink ports

Uplink port and downlink port are two port roles in monitor link groups:

- Uplink ports are the monitored ports. The state of a monitor link group adapts to that of its member uplink ports. When a monitor link group contains no uplink port or when all the uplink ports are down, the monitor link group becomes down. As long as one member uplink port is up, the monitor link group stays up.

- Downlink ports are the monitoring ports. The state of the downlink ports in a monitor link group adapts to that of the monitor link group. When the state of a monitor link group changes, the state of its member downlink ports change accordingly. The state of the downlink ports in a monitor link group is always consistent with that of the monitor link group.

## Uplink/Downlink

The uplink is the link that connects the uplink ports in a monitor link group, and the downlink is the link that connects the downlink ports.

## How Monitor Link works

A monitor link group works independently of other monitor link groups. When a monitor link group contains no uplink port or when all its uplink ports are down, the monitor link group goes down and forces all downlink ports down at the same time. When any uplink port goes up, the monitor link group goes up and brings up all the downlink ports.

HP does not recommend manually shutting down or bringing up the downlink ports in a monitor link group.

# Configuring Monitor Link

## Configuration prerequisites

Make sure that the port is not the member port of any aggregation group or service loopback group.

## Creating a monitor link group

Step	Command
1. Enter system view.	<b>system-view</b>
2. Create a monitor link group, and enter monitor link group view.	<b>monitor-link group</b> <i>group-id</i>

## Configuring monitor link group member ports

You can configure member ports for a monitor link group either in monitor link group view or interface view. The configurations made in these two views lead to the same result.

You can assign a Layer 2 Ethernet port or Layer 2 aggregate interface to a monitor link group as a member port.

A port can be assigned to only one monitor link group.

Configure uplink ports prior to downlink ports to avoid undesired down/up state changes on the downlink ports.

### In monitor link group view

To configure member ports for a monitor link group in monitor link group view:

Step	Command
1. Enter system view.	<b>system-view</b>
2. Enter monitor link group view.	<b>monitor-link group</b> <i>group-id</i>
3. Configure member ports for the monitor link group.	<b>port</b> <i>interface-type interface-number</i> { <b>uplink</b>   <b>downlink</b> }

### In interface view

To configure member ports for a monitor link group in interface view:

Step	Command
1. Enter system view.	<b>system-view</b>
2. Enter Layer 2 Ethernet interface view or Layer 2 aggregate interface view.	<b>interface</b> <i>interface-type interface-number</i>
3. Configure the current interface as a member of a monitor link group.	<b>port monitor-link group</b> <i>group-id</i> { <b>uplink</b>   <b>downlink</b> }

## Displaying and maintaining Monitor Link

Task	Command	Remarks
Display monitor link group information.	<b>display monitor-link group</b> { <i>group-id</i>   <b>all</b> } [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]	Available in any view

## Monitor Link configuration example

### Network requirements

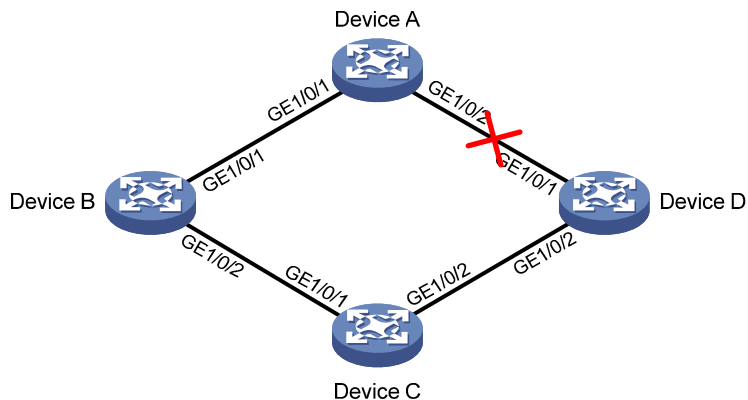
As shown in [Figure 29](#), Device C is a smart link device, and Device A, Device B, and Device D are associated devices. Traffic of VLANs 1 through 30 on Device C is dual-uplinked to Device A through a smart link group.

Implement dual uplink backup on Device C, and make sure that when the link between Device A and Device B (or Device D) fails, Device C can sense the link fault and perform uplink switchover in the smart link group.

For more information about Smart Link, see "[Configuring Smart Link](#)."



Figure 29 Network diagram



## Configuration procedure

### 1. Configure Device C:

# Create VLANs 1 through 30, map these VLANs to MSTI 1, and activate MST region configuration.

```
<DeviceC> system-view
[DeviceC] vlan 1 to 30
[DeviceC] stp region-configuration
[DeviceC-mst-region] instance 1 vlan 1 to 30
[DeviceC-mst-region] active region-configuration
[DeviceC-mst-region] quit
```

# Disable the spanning tree feature on GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 separately, configure them as trunk ports, and assign them to VLANs 1 through 30.

```
[DeviceC] interface gigabitethernet 1/0/1
[DeviceC-GigabitEthernet1/0/1] undo stp enable
[DeviceC-GigabitEthernet1/0/1] port link-type trunk
[DeviceC-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceC-GigabitEthernet1/0/1] quit
[DeviceC] interface gigabitethernet 1/0/2
[DeviceC-GigabitEthernet1/0/2] undo stp enable
[DeviceC-GigabitEthernet1/0/2] port link-type trunk
[DeviceC-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceC-GigabitEthernet1/0/2] quit
```

# Create smart link group 1, and configure all the VLANs mapped to MSTI 1 as the protected VLANs for smart link group 1.

```
[DeviceC] smart-link group 1
[DeviceC-smlk-group1] protected-vlan reference-instance 1
```

# Configure GigabitEthernet 1/0/1 as the master port and GigabitEthernet 1/0/2 as the slave port for smart link group 1.

```
[DeviceC-smlk-group1] port gigabitethernet 1/0/1 master
[DeviceC-smlk-group1] port gigabitethernet 1/0/2 slave
```

# Enable the smart link group to transmit flush messages.

```
[DeviceC-smlk-group1] flush enable
[DeviceC-smlk-group1] quit
```

## 2. Configure Device A:

# Create VLANs 1 through 30.

```
<DeviceA> system-view
```

```
[DeviceA] vlan 1 to 30
```

# Configure GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 as trunk ports, assign them to VLANs 1 through 30, and enable flush message receiving on them.

```
[DeviceA] interface gigabitethernet 1/0/1
```

```
[DeviceA-GigabitEthernet1/0/1] port link-type trunk
```

```
[DeviceA-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
```

```
[DeviceA-GigabitEthernet1/0/1] smart-link flush enable
```

```
[DeviceA-GigabitEthernet1/0/1] quit
```

```
[DeviceA] interface gigabitethernet 1/0/2
```

```
[DeviceA-GigabitEthernet1/0/2] port link-type trunk
```

```
[DeviceA-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
```

```
[DeviceA-GigabitEthernet1/0/2] smart-link flush enable
```

```
[DeviceA-GigabitEthernet1/0/2] quit
```

## 3. Configure Device B:

# Create VLANs 1 through 30.

```
<DeviceB> system-view
```

```
[DeviceB] vlan 1 to 30
```

# Configure GigabitEthernet 1/0/1 as a trunk port, assign it to VLANs 1 through 30, and enable flush message receiving on it.

```
[DeviceB] interface gigabitethernet 1/0/1
```

```
[DeviceB-GigabitEthernet1/0/1] port link-type trunk
```

```
[DeviceB-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
```

```
[DeviceB-GigabitEthernet1/0/1] smart-link flush enable
```

```
[DeviceB-GigabitEthernet1/0/1] quit
```

# Configure GigabitEthernet 1/0/2 as a trunk port, assign it to VLANs 1 through 30, disable the spanning tree feature, and enable flush message receiving on it.

```
[DeviceB] interface gigabitethernet 1/0/2
```

```
[DeviceB-GigabitEthernet1/0/2] port link-type trunk
```

```
[DeviceB-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
```

```
[DeviceB-GigabitEthernet1/0/2] undo stp enable
```

```
[DeviceB-GigabitEthernet1/0/2] smart-link flush enable
```

```
[DeviceB-GigabitEthernet1/0/2] quit
```

# Create monitor link group 1, and then configure GigabitEthernet 1/0/1 as an uplink port and GigabitEthernet 1/0/2 as a downlink port for monitor link group 1.

```
[DeviceB] monitor-link group 1
```

```
[DeviceB-mtlk-group1] port gigabitethernet 1/0/1 uplink
```

```
[DeviceB-mtlk-group1] port gigabitethernet 1/0/2 downlink
```

```
[DeviceB-mtlk-group1] quit
```

## 4. Configure Device D:

# Create VLANs 1 through 30.

```
<DeviceD> system-view
```

```
[DeviceD] vlan 1 to 30
```

# Configure GigabitEthernet 1/0/1 as a trunk port, assign it to VLANs 1 through 30, and enable flush message receiving on it.

```
[DeviceD] interface gigabitethernet 1/0/1
[DeviceD-GigabitEthernet1/0/1] port link-type trunk
[DeviceD-GigabitEthernet1/0/1] port trunk permit vlan 1 to 30
[DeviceD-GigabitEthernet1/0/1] smart-link flush enable
[DeviceD-GigabitEthernet1/0/1] quit
```

# Configure GigabitEthernet 1/0/2 as a trunk port, assign it to VLANs 1 through 30, disable the spanning tree feature, and enable flush message receiving on it.

```
[DeviceD] interface gigabitethernet 1/0/2
[DeviceD-GigabitEthernet1/0/2] port link-type trunk
[DeviceD-GigabitEthernet1/0/2] port trunk permit vlan 1 to 30
[DeviceD-GigabitEthernet1/0/2] undo stp enable
[DeviceD-GigabitEthernet1/0/2] smart-link flush enable
[DeviceD-GigabitEthernet1/0/2] quit
```

# Create monitor link group 1, and then configure GigabitEthernet 1/0/1 as an uplink port and GigabitEthernet 1/0/2 as a downlink port for monitor link group 1.

```
[DeviceD] monitor-link group 1
[DeviceD-mtlk-group1] port gigabitethernet 1/0/1 uplink
[DeviceD-mtlk-group1] port gigabitethernet 1/0/2 downlink
[DeviceD-mtlk-group1] quit
```

#### 5. Verify the configuration:

Use the **display monitor-link group** command to display the monitor link group information on devices. For example, when GigabitEthernet 1/0/2 on Device A goes down due to a link fault:

# Display information about monitor link group 1 on Device B.

```
[DeviceB] display monitor-link group 1
Monitor link group 1 information:
Group status: UP
Last-up-time: 16:37:20 2009/4/21
Last-down-time: 16:35:26 2009/4/21
Member Role Status

GigabitEthernet1/0/1 UPLINK UP
GigabitEthernet1/0/2 DOWNLINK UP
```

# Display information about monitor link group 1 on Device D.

```
[DeviceD] display monitor-link group 1
Monitor link group 1 information:
Group status: DOWN
Last-up-time: 16:35:27 2009/4/21
Last-down-time: 16:37:19 2009/4/21
Member Role Status

GigabitEthernet1/0/1 UPLINK DOWN
GigabitEthernet1/0/2 DOWNLINK DOWN
```

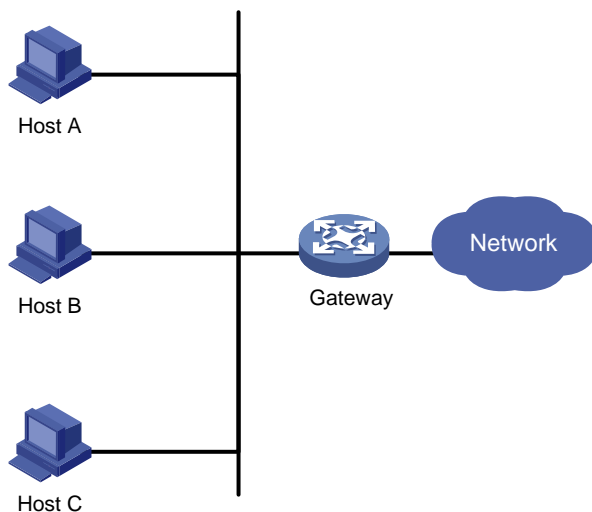
# Configuring VRRP (available only on the HP 5500 EI)

- The term *router* in this document refers to both routers and Layer 3 switches.
- You can perform interface-specific VRRP configuration only on Layer 3 Ethernet interfaces, VLAN interfaces, and Layer 3 aggregate interfaces, unless otherwise specified. You can set an Ethernet port as a Layer 3 interface by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).
- VRRP cannot be configured on interfaces in aggregation groups.

## VRRP overview

Typically, as shown in [Figure 30](#), you can configure a default route with the gateway as the next hop for every host on a network segment. All packets destined to other network segments are sent over the default route to the gateway, which then forwards the packets. However, when the gateway fails, all the hosts that use the gateway as the default next-hop router fail to communicate with external networks.

**Figure 30 LAN networking**



Configuring a default route for network hosts facilitates your configuration, but also requires high performance stability of the device that acts as the gateway. Using more egress gateways is a common way to improve system reliability, but introduces the problem of routing among the egresses.

Virtual Router Redundancy Protocol (VRRP) is designed to address this problem. VRRP adds a group of routers that can act as network gateways to a VRRP group, which forms a virtual router. Routers in the VRRP group elect a master through the VRRP election mechanism to act as a gateway, and hosts on a LAN only need to configure the virtual router as their default network gateway.

VRRP is an error-tolerant protocol, which improves the network reliability and simplifies configurations on hosts. On a multicast and broadcast LAN such as Ethernet, VRRP provides highly reliable default links without configuration changes (such as dynamic routing protocols, route discovery protocols) when a router fails, and prevent network interruption because of a single link failure.

VRRP operates in either of the following modes:

- **Standard protocol mode**—Includes two versions VRRPv2 and VRRPv3 based on RFCs. VRRPv2 is based on IPv4, and VRRPv3 is based on IPv6. The two versions implement the same functions but are applied in different network environments. For more information, see "[VRRP standard protocol mode](#)."
- **Load balancing mode**—Extends the standard protocol mode and realizes load balancing. For more information, see "[VRRP load balancing mode](#)."

## VRRP standard protocol mode

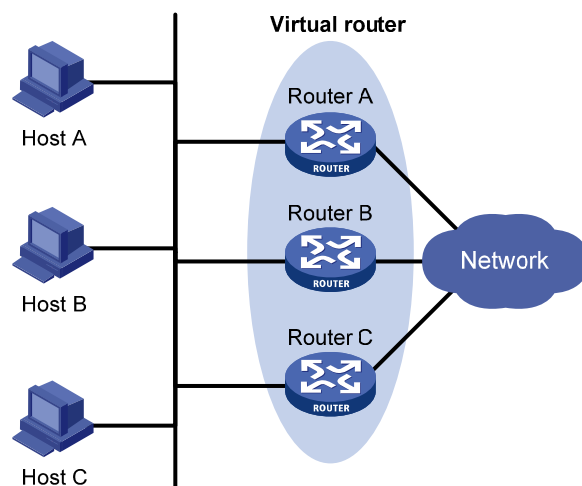
### Introduction to VRRP group

VRRP combines a group of routers (including a master and multiple backups) on a LAN into a virtual router called VRRP group.

A VRRP group has the following features:

- A virtual router has a virtual IP address. A host on the LAN only needs to know the IP address of the virtual router and uses the IP address as the next hop of the default route.
- Every host on the LAN communicates with external networks through the virtual router.
- Routers in the VRRP group elect a master that acts as the gateway according to their priorities. The other routers function as the backups. When the master fails, to make sure that the hosts in the network segment can uninterruptedly communicate with the external networks, the backups in the VRRP group elect a new gateway to undertake the responsibility of the failed master.

**Figure 31 Network diagram**



As shown in [Figure 31](#), Router A, Router B, and Router C form a virtual router, which has its own IP address. Hosts on the Ethernet use the virtual router as the default gateway.

The router with the highest priority among the three routers is elected as the master to act as the gateway, and the other two are backups.

The IP address of the virtual router can be either an unused IP address on the segment where the VRRP group resides or the IP address of an interface on a router in the VRRP group. In the latter case, the router is called the IP address owner.

Only one IP address owner can be configured for a VRRP group.

A router in a VRRP group can be in master, backup, or initialize status.

## VRRP priority

VRRP determines the role (master or backup) of each router in a VRRP group by priority. A router with a higher priority is more likely to become the master.

VRRP priority is in the range of 0 to 255. The greater the number, the higher the priority. Priorities 1 to 254 are configurable. Priority 0 is reserved for special uses and priority 255 for the IP address owner. When a router acts as the IP address owner, its running priority is always 255. That is, the IP address owner in a VRRP group acts as the master as long as it works properly.

## Operation mode

A router in a VRRP group operates in either of the following modes:

- **Non-preemptive mode**—When a router in the VRRP group becomes the master, it stays as the master as long as it operates normally, even if a backup is assigned a higher priority later.
- **Preemptive mode**—When a backup finds its priority higher than that of the master, the backup sends VRRP advertisements to start a new master election in the VRRP group and becomes the master. Accordingly, the original master becomes a backup.

## Authentication mode

To avoid attacks from unauthorized users, VRRP adds authentication keys into packets for authentication. VRRP provides the following authentication modes:

- **simple**—Simple text authentication

A router sending a packet fills an authentication key into the packet, and the router receiving the packet compares its local authentication key with that of the received packet. If the two authentication keys are the same, the received VRRP packet is considered legitimate. Otherwise, the received packet is considered invalid.

- **md5**—MD5 authentication

A router computes the digest of a packet to be sent by using the authentication key and MD5 algorithm and saves the result in the authentication header. The router that receives the packet performs the same operation by using the authentication key and MD5 algorithm, and compares the result with the content in the authentication header. If the results are the same, the router that receives the packet considers the packet an authentic and valid VRRP packet. Otherwise, the router considers the packet invalid.

On a secure network, you can choose not to set the authentication mode.

## VRRP timers

VRRP timers include VRRP advertisement interval timer and VRRP preemption delay timer.

### VRRP advertisement interval timer

The master in a VRRP group periodically sends VRRP advertisements to inform the other routers in the VRRP group that it operates properly.

You can adjust the interval for sending VRRP advertisements by setting the VRRP advertisement interval timer. If a backup receives no advertisements in a period three times the interval, the backup regards itself as the master and sends VRRP advertisements to start a new master election.

## VRRP preemption delay timer

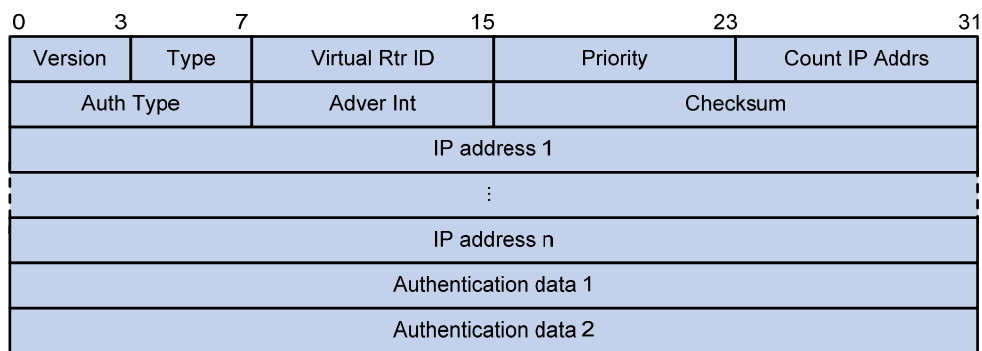
To avoid frequent state changes among members in a VRRP group and provide the backups enough time to collect information (such as routing information), each backup waits for a period of time (the preemption delay time) after it receives an advertisement with the priority lower than the local priority, then sends VRRP advertisements to start a new master election in the VRRP group and becomes the master.

## Packet format

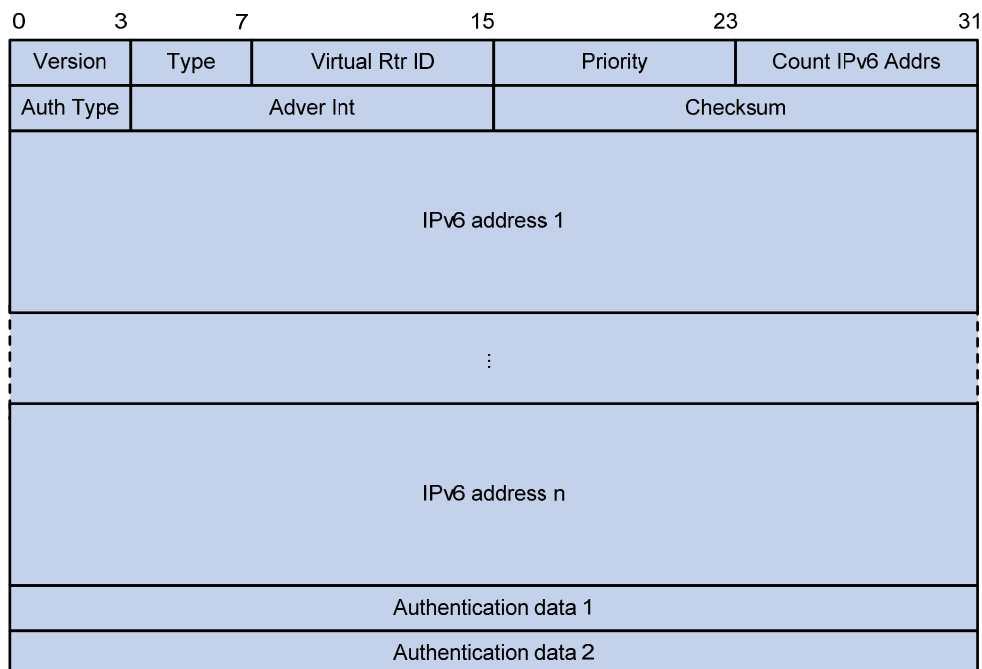
The master multicasts VRRP packets periodically to declare its existence. VRRP packets are also used for checking the parameters of the virtual router and electing the master.

VRRP packets are encapsulated in IP packets, with the protocol number being 112. [Figure 32](#) shows the format of a VRRPv2 packet and [Figure 33](#) shows the format of a VRRPv3 packet.

**Figure 32 Format of a VRRPv2 packet**



**Figure 33 Format of a VRRPv3 packet**



A VRRP packet comprises the following fields:

- **Version**—Version number of the protocol, 2 for VRRPv2 and 3 for VRRPv3.
- **Type**—Type of the VRRPv2 or VRRPv3 packet. Only one VRRP packet type is present, that is, VRRP advertisement, which is represented by 1.
- **Virtual Rtr ID (VRID)**—ID of the virtual router, that is, ID of the VRRP group. It ranges from 1 to 255.
- **Priority**—Priority of the router in the VRRP group, in the range of 0 to 255. A greater value represents a higher priority.
- **Count IP Adrs/Count IPv6 Adrs**—Number of virtual IPv4 or IPv6 addresses for the VRRP group. A VRRP group can have multiple virtual IPv4 or IPv6 addresses.
- **Auth Type**—Authentication type. 0 means no authentication, 1 means simple text authentication, and 2 means MD5 authentication. VRRPv3 does not support MD5 authentication.
- **Adver Int**—Interval for sending advertisement packets. For VRRPv2, the interval is in seconds and defaults to 1. For VRRPv3, the interval is in centiseconds and defaults to 100.
- **Checksum**—16-bit checksum for validating the data in VRRP packets.
- **IP Address/IPv6 Address**—Virtual IPv4 or IPv6 address entry of the VRRP group. The Count IP Adrs or Count IPv6 Adrs field defines the number of the virtual IP v4 or IPv6 addresses.
- **Authentication Data**—Authentication key. This field is used only for simple authentication and is 0 for any other authentication modes.

## Principles of VRRP

- Routers in a VRRP group determine their roles by priority. The router with the highest priority is the master, and the others are the backups. The master periodically sends VRRP advertisements to notify the backups that it is operating properly, and each of the backups starts a timer to wait for advertisements from the master.
- In preemptive mode, when a backup receives a VRRP advertisement, it compares the priority in the packet with its own priority. If the priority of the backup is higher, the backup becomes the master. Otherwise, it remains as a backup. With the preemptive mode, a VRRP group always has a router with the highest priority as the master for packet forwarding.
- In non-preemptive mode, a router in the VRRP group remains as a master or backup as long as the master does not fail. A backup does not become the master even if it is configured with a higher priority. The non-preemptive mode helps avoid frequent switchover between the master and backups.
- If the timer of a backup expires but the backup still does not receive any VRRP advertisement, it considers that the master fails. In this case, the backup considers itself as the master and sends VRRP advertisements to start a new master election.

The VRRP group configuration might be different on routers, and network problems might exist, so multiple master routers might exist in one VRRP group. These master routers will elect one master according to their priorities and IP addresses. The router with the highest priority wins the election. If a tie exists in the priority, the router with the highest IP address wins.

After a backup router receives an advertisement, it compares its priority against that carried in the advertisement. If its priority is higher than that carried in the advertisement, it takes over the master.

## VRRP tracking

To enable the VRRP tracking function, configure the routers in the VRRP group to operate in preemptive mode first, so that only the router with the highest priority can always operate as the master for packet forwarding.



## Tracking a specified interface

The interface tracking function expands the backup functionality of VRRP. It provides backup not only when the interface to which a VRRP group is assigned fails but also when other interfaces (such as uplink interfaces) on the router become unavailable.

If the uplink interface of a router in a VRRP group fails, usually the VRRP group cannot be aware of the uplink interface failure. If the router is the master of the VRRP group, hosts on the LAN are not able to access external networks because of the uplink failure. This problem can be solved by tracking a specified uplink interface. If the tracked uplink interface is down or removed, the priority of the master is automatically decreased by a specified value and a higher priority router in the VRRP group becomes the master.

## Monitoring a track entry

By monitoring a track entry, you can:

- Monitor an uplink and change the priority of the router according to the state of the uplink. If the uplink fails, hosts in the LAN cannot access external networks through the router. In this case, the state of the monitored track entry is negative and the priority of the router decreases by a specified value. Then, a higher priority router in the VRRP group becomes the master to maintain the proper communication between the hosts in the LAN and external networks.
- Monitor the master on a backup. When the master fails, the backup immediately preempts as the master to maintain normal communication.

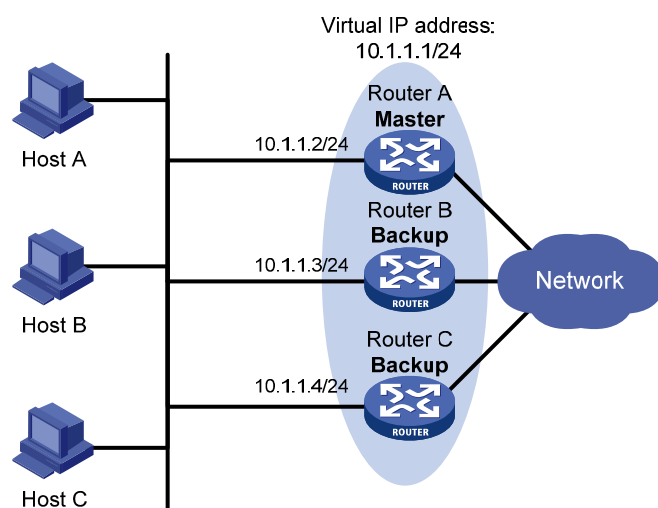
For more information about track entries, see "[Configuring track](#)."

# VRRP application (taking IPv4-based VRRP for example)

## Master/backup

In master/backup mode, only the master forwards packets. When the master fails, a new master is elected from the backups. This mode requires only one VRRP group, in which each router holds a different priority and the one with the highest priority becomes the master, as shown in [Figure 34](#).

**Figure 34 VRRP in master/backup mode**



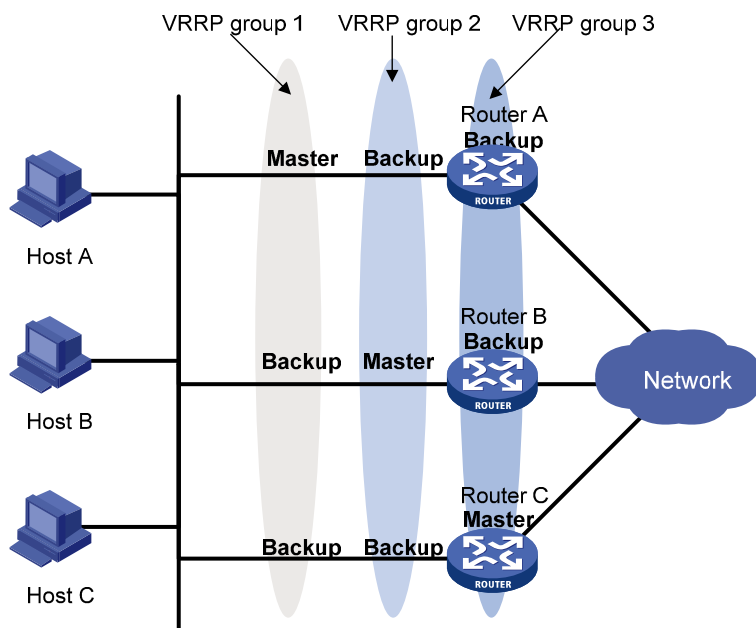
Assume that Router A is the master and therefore can forward packets to external networks, whereas Router B and Router C are backups and are thus in the state of listening. If Router A fails, Router B and Router C elect for a new master to forward packets to hosts on the LAN.

## Load sharing

More than one VRRP group can be created on an interface of a router to allow the router to be the master of one VRRP group but a backup of another at the same time.

In load sharing mode, multiple routers provide services simultaneously. This mode requires two or more VRRP groups, each of which comprises a master and one or more backups. The masters of the VRRP groups are assumed by different routers, as shown in Figure 35.

Figure 35 VRRP in load sharing mode



A router can be in multiple VRRP groups and hold a different priority in a different group.

As shown in Figure 35, the following VRRP groups are present:

- **VRRP group 1**—Router A is the master; Router B and Router C are the backups.
- **VRRP group 2**—Router B is the master; Router A and Router C are the backups.
- **VRRP group 3**—Router C is the master; Router A and Router B are the backups.

For load sharing among Router A, Router B, and Router C, hosts on the LAN need to be configured to use VRRP group 1, 2, and 3 as the default gateways. When you configure VRRP priorities, make sure that each router holds such a priority in each VRRP group that it will take the expected role in the group.

## VRRP load balancing mode

### Overview

When VRRP is operating in standard protocol mode, only the master can forward packets and the backups are in the state of listening. You can create multiple VRRP groups to share the load among multiple routers, but hosts on the LAN need to be configured with different gateways, thus making the configuration complicated.

In load balancing mode, VRRP provides load balancing in addition to virtual gateway redundancy by mapping a virtual IP address to multiple virtual MAC addresses to assign each router in a VRRP group

one virtual MAC address. In this way, each router in this VRRP group can respond to ARP requests (in an IPv4 network) or ND requests (in an IPv6 network) from corresponding hosts, so that different hosts can send packets to different routers, and each router in the VRRP group can forward packets. In load balancing mode, you need to create only one VRRP group to balance load among multiple routers, instead of allowing one router to bear the load but other routers to stay idle.

VRRP load balancing mode is based on VRRP standard protocol mode, so mechanisms, such as master election, preemption, and tracking functions, in the standard protocol mode are also supported in the load balancing mode. In addition, VRRP load balancing mode has new mechanisms, which are introduced in the following sections.

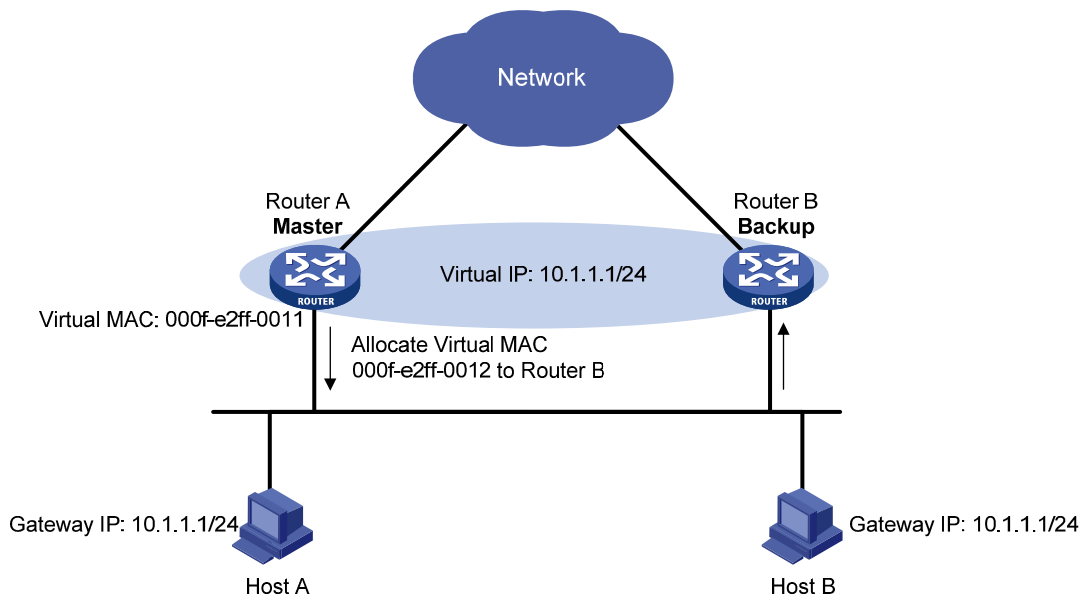
## Assigning virtual MAC addresses

When VRRP is operating in load balancing mode, the master assigns virtual MAC addresses to the routers in the VRRP group and answers the ARP requests or ND requests from different hosts. The backup routers, however, do not answer the ARP requests or ND requests from the hosts.

Assume that a VRRP group is operating in an IPv4 network. The following describes how the load balancing mode works:

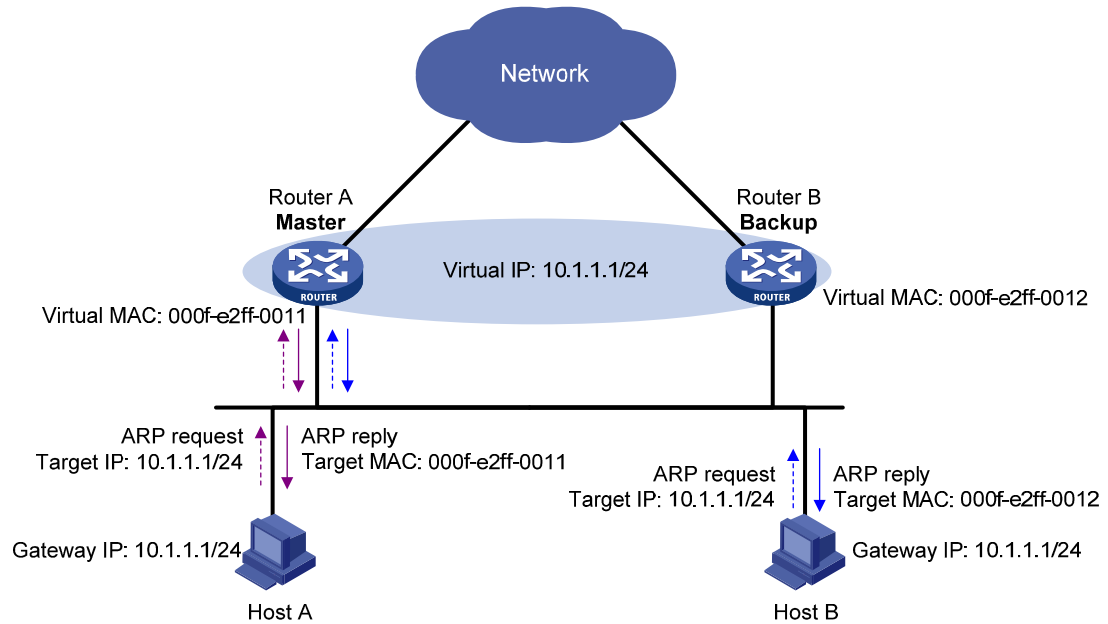
1. The master assigns virtual MAC addresses to the routers (including the master itself and the backups) in the VRRP group. For example, as shown in [Figure 36](#), the virtual IP address of the VRRP group is 10.1.1.1/24; Router A is the master; Router B and Router C are the backups. Router A assigns 000f-e2ff-0011 to itself, and 000f-e2ff-0012 to Router B.

**Figure 36 Allocating virtual MAC addresses**



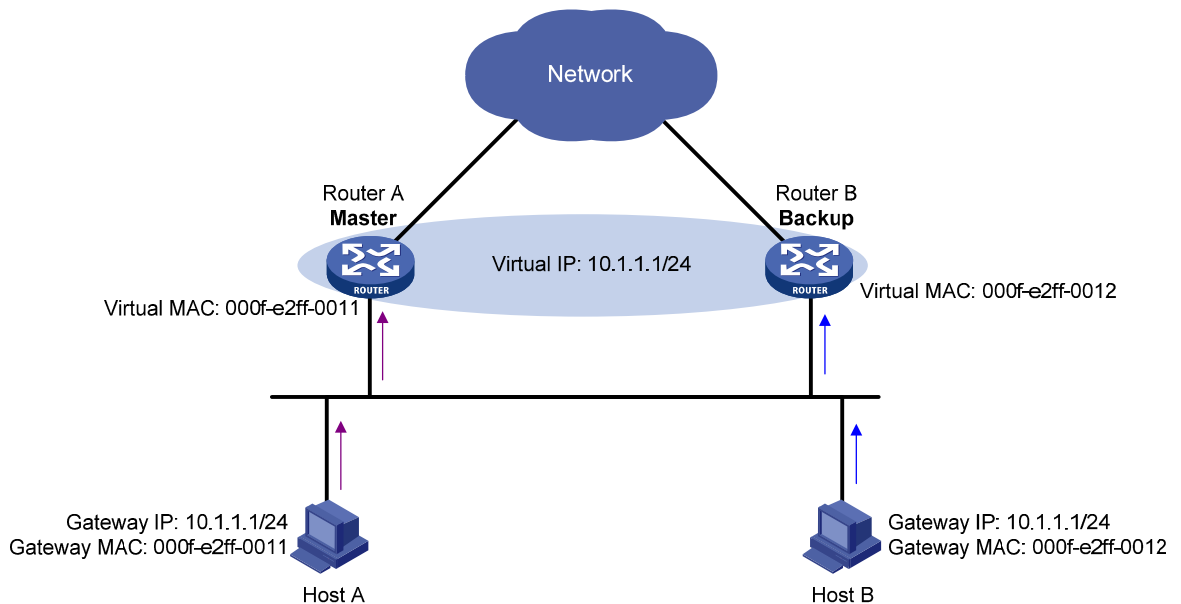
2. After receiving an ARP request destined for the virtual IP address of the VRRP group from a host, the master, based on the load balancing algorithm, uses a corresponding virtual MAC address to answer the ARP request. For example, as shown [Figure 37](#), when Host A sends an ARP request to retrieve the MAC address of gateway 10.1.1.1, the master (Router A), after receiving the request, returns the virtual MAC address of Router A to Host A; when Host B sends an ARP request to retrieve the MAC address of gateway 10.1.1.1, the master (Router A), after receiving the request, returns the virtual MAC address of Router B to Host B.

Figure 37 Answering ARP requests



3. Different hosts send packets to different routers according to the requested virtual MAC addresses. For example, as shown in Figure 38, Host A regards the virtual MAC address of Router A as the gateway MAC address, so it sends packets to Router A for forwarding; Host B regards the virtual MAC address of Router B as the gateway MAC address, so it sends packets to Router B for forwarding.

Figure 38 Sending packets to different routers for forwarding



# Virtual forwarder

## Creating a virtual forwarder

Virtual MAC addresses help different hosts transmit packets to different routers in a VRRP group. To enable the routers in the VRRP group to forward the packets, be sure to create virtual forwarders (VFs) on the routers. Each VF associates with a virtual MAC address in the VRRP group and forwards packets destined to this virtual MAC address.

The following describes how VFs are created on the routers in a VRRP group:

1. The master assigns virtual MAC addresses to all routers in the VRRP group. After learning its virtual MAC address, a router in the VRRP group creates a VF that corresponds to this MAC address, and becomes the owner of this VF.
2. The router advertises the VF information to the other routers in the VRRP group.
3. After receiving the VF advertisement, each of the other routers creates the advertised VF.

As described in the preceding steps, each router in the VRRP group creates not only a VF corresponding to its virtual MAC address, but also VFs advertised by the other routes in the VRRP group..

## VF weight and priority

The weight of a VF indicates the forwarding capability of a router. A higher weight indicates a higher forwarding capability. When the weight is lower than the lower limit of failure, the router cannot be capable of forwarding packets for the hosts.

The priority of a VF determines the VF state. Among the VFs that correspond to the same virtual MAC address on different routers in the VRRP group, the VF with the highest priority is in the active state and is known as the active virtual forwarder (AVF), which forwards packets; other VFs are in the listening state and are known as the listening virtual forwarders (LVFs), which monitor the state of the AVF. The priority value of a VF ranges from 0 to 255, where 255 is reserved for the VF owner. If the weight of a VF owner is higher than or equal to the lower limit of failure, the priority value of the VF owner is 255.

The priority value of a VF is calculated based on its weight:

- If the weight of a VF is higher than or equal to the lower limit of failure, and the router where the VF resides is the owner of the VF, the priority value of the VF is 255.
- If the weight of a VF is higher than or equal to the lower limit of failure, but the router where the VF resides is not the owner of the VF, the priority value of the VF is  $\text{weight}/(\text{number of local AVFs} + 1)$
- If the weight of a VF is lower than the lower limit of failure, the priority value of the VF is 0.

## VF backup

The VFs corresponding to a virtual MAC address on different routers in the VRRP group back up one another.

Figure 39 VF information

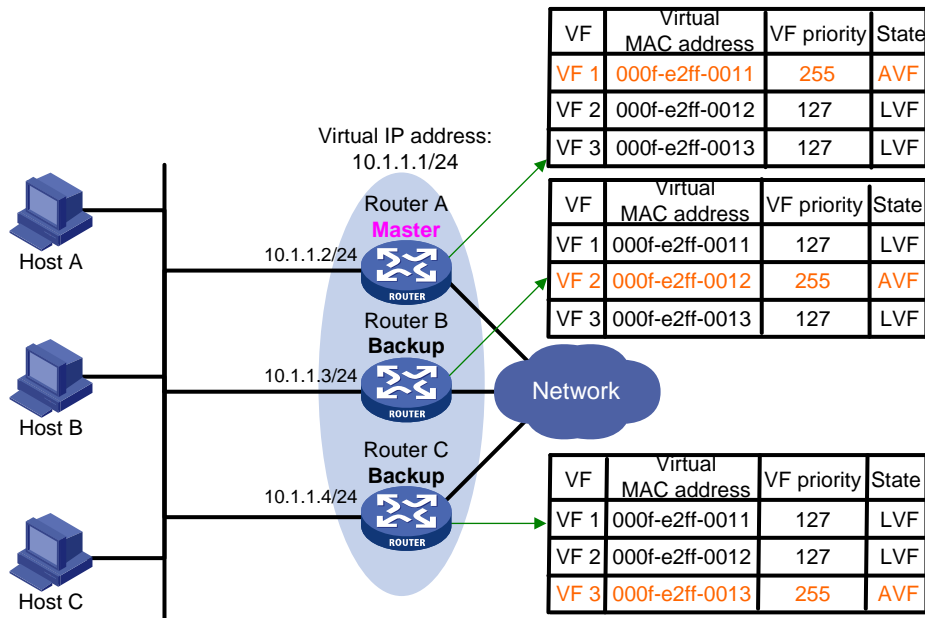


Figure 39 illustrates the VF information on each router in the VRRP group and how the routers back up one another. The master, Router A, assigns virtual MAC addresses 000f-e2ff-0011, 000f-e2ff-0012, and 000f-e2ff-0013 to itself, Router B, and Router C, respectively. The VFs corresponding to these three virtual MAC addresses, VF 1, VF 2, and VF 3, are created on each of the three routers, and the VFs corresponding to the same virtual MAC address on different routers back up one another. For example, VF 1 on Router A, Router B, and Router C can implement backup.

- Router A is the owner of VF 1, and the priority value of VF 1 on Router A is 255. In this case, VF 1 on Router A acts as the AVF to forward the packets destined for virtual MAC address 000f-e2ff-0011.
- The priority value of VF 1 on Router B and Router C is  $\text{weight}/(\text{number of local AVFs} + 1)$ , that is,  $255/(1 + 1) = 127$ , which is lower than that of VF 1 on Router A. In this case, VF 1 on both Router B and Router C acts as the LVF to listen to the status of VF 1 on Router A.
- When VF 1 on Router A fails, VF 1 on Router B and Router C elects the one with a higher priority value as the new AVF, responsible for forwarding the packets destined for virtual MAC address 000f-e2ff-0011.

**NOTE:**

A VF always operates in preemptive mode. When an LVF finds its priority value higher than that in the advertisement sent by the AVF, the LVF declares itself as the AVF.

**VF timers**

When the AVF on a router fails, the newly elected AVF on another router creates a redirect timer and a timeout timer for the failed AVF.

- **Redirect Timer**—Before this timer times out, the master still uses the virtual MAC address corresponding to the failed AVF to respond to ARP/ND requests from the hosts, and the VF owner can share traffic load if the VF owner resumes normal operation within this time. When this timer times out, the master stops using the virtual MAC address corresponding to the failed AVF to respond to ARP/ND requests from the hosts.

- **Timeout Timer**—The duration that the new AVF takes over the VF owner. Before this timer times out, all the routers in the VRRP group keep the failed AVF, and the new AVF forwards the packets destined for the virtual MAC address corresponding to the failed AVF. When this timer times out, all the routers in the VRRP group remove the failed AVF, and the new AVF stops forwarding the packets destined for the virtual MAC address corresponding to the failed AVF.

## VF tracking

The AVF forwards packets destined to the MAC address of the AVF. If the uplink of the AVF fails and no LVF is notified to take over the AVF's work, hosts (on the LAN) that use the MAC address of the AVF as their gateway MAC address cannot access the external network.

This problem can be solved by the VF tracking function. You can monitor the uplink state by using network quality analyzer (NQA) and bidirectional forwarding detection (BFD), and establish the collaboration between the VF and the NQA or between the VF and the BFD through the tracking function. When the uplink fails, the state of the monitored track entry changes to negative and the weight of the VF decreases by a specified value. Then, the VF with a higher priority becomes the AVF and forwards packets.

The VF tracking function can also work on an LVF to monitor its corresponding AVF on another router. When the AVF fails, the LVF immediately takes over the AVF to ensure uninterrupted network communications.

## Packet types

VRRP standard protocol mode defines only VRRP advertisement. Only the master in a VRRP group periodically sends VRRP advertisements, and the backups do not send VRRP advertisements.

VRRP load balancing mode defines the following types of packets:

- **Advertisement**—VRRP advertises VRRP group state and information about the VF that is in the active state. Both the master and the backups periodically send VRRP advertisements.
- **Request**—If a backup is not the VF owner, it sends a request to ask the master to assign a virtual MAC address.
- **Reply**—When receiving a request, the master sends a reply to the backup router to assign a virtual MAC address. After receiving the reply, the backup router creates a VF that corresponds to the virtual MAC address, and then becomes the owner of this VF.
- **Release**—When a VF owner fails, the router that takes over its responsibility sends a release after a specified period of time to notify the other routers in the VRRP group to delete the VF of the failed VF owner.

---

### NOTE:

The format of these packets is similar to that of the advertisement in VRRP standard protocol mode except that a packet used in load balancing mode is appended with option field, which contains information for load balancing.

---

# Configuring VRRP for IPv4

## VRRP for IPv4 configuration task list

To form a VRRP group, perform the following configurations on each device in the VRRP group.

Complete these tasks to configure VRRP for IPv4:

Task	Remarks
Configuring a VRRP operation mode	Optional
Specifying the type of MAC addresses mapped to virtual IP addresses	Optional When VRRP is operating in load balancing mode, this configuration is not effective.
Creating a VRRP group and configuring virtual IP address	Required
Configuring router priority, preemptive mode and tracking function	Optional
Configuring VF tracking	Optional The VF tracking function is effective only when VRRP is operating in load balancing mode.
Configuring VRRP packet attributes	Optional
Enabling the trap function for VRRP	Optional

## Configuring a VRRP operation mode

VRRP can operate in either of the following modes:

- **Standard protocol mode**—When VRRP is operating in this mode, only the master in a VRRP group is responsible for forwarding packets.
- **Load balancing mode**—When VRRP is operating in this mode, all the routers (master and backups) that have the AVF in a VRRP group can forward packets, thus implementing load balancing.

After the VRRP operation mode is specified on a router, all VRRP groups on the router operate in the specified operation mode.

To configure a VRRP operation mode:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Configure a VRRP operation mode.	Configure VRRP to operate in standard protocol mode: <b>undo vrrp mode</b> Configure VRRP to operate in load balancing mode: <b>vrrp mode load-balance</b>	Use either command. By default, VRRP operates in standard protocol mode.

## Specifying the type of MAC addresses mapped to virtual IP addresses

After you specify the type of MAC addresses mapped to the virtual IP addresses of VRRP groups and create a VRRP group, the master in the VRRP group uses the specified type of MAC address as the source MAC address for sending packets and uses the specified type of MAC address to answer ARP requests



from hosts so that the hosts in the internal network can learn the mapping between the IP address and the MAC address.

The following types of MAC addresses are available to be mapped to the virtual IP address of a VRRP group:

- **Virtual MAC to virtual IP mapping**—By default, a virtual MAC address is automatically created for a VRRP group when the VRRP group is created, and the virtual IP address of the VRRP group is mapped to the virtual MAC address. When such a mapping is adopted, the hosts in the internal network do not need to update the mapping between the IP address and MAC address when the master changes.
- **Real MAC to virtual IP mapping**—In case that an IP address owner exists in a VRRP group, if the virtual IP address is mapped to the virtual MAC address, two MAC addresses are mapped to one IP address. To avoid such a problem, map the virtual IP address of the VRRP group to the real MAC address of an interface to forward the packets from a host to the IP address owner.

### Configuration guidelines

- When VRRP is operating in load balancing mode, a virtual IP address is always mapped to a virtual MAC address regardless of which type of MAC addresses are specified to be mapped to virtual IP addresses.
- Specify the type of the MAC addresses mapped to the virtual IP addresses before creating a VRRP group. Otherwise, you cannot change the type of the MAC addresses mapped to virtual IP addresses.
- If VRRP groups with the same ID are created on multiple interfaces of a device, and the VRRP advertisements of these VRRP groups are to be sent through QinQ networks, HP recommends you to map the real MAC addresses of the interfaces to the virtual IP addresses of these VRRP groups. Otherwise, the VRRP advertisements of these VRRP groups cannot be sent successfully.

### Configuration procedure

To specify the type of MAC addresses mapped to virtual IP addresses:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Specify the type of MAC addresses mapped to virtual IP addresses.	<b>vrrp method { real-mac   virtual-mac }</b>	Optional. Virtual MAC address by default.

## Creating a VRRP group and configuring virtual IP address

When creating a VRRP group on an interface, configure a virtual IP address for the VRRP group. If the interface connects to multiple sub-networks, you can configure multiple virtual IP addresses for the VRRP group to realize router backup on different sub-networks.

A VRRP group is automatically created when you specify the first virtual IP address for the VRRP group. If you specify another virtual IP address for the VRRP group later, the virtual IP address is added to the virtual IP address list of the VRRP group.

Do not create VRRP groups on the VLAN interface of a super VLAN. Otherwise, network performance might be affected.

## Configuration guidelines

- When VRRP is operating in standard protocol mode, the virtual IP address of a VRRP group can be either an unused IP address on the segment where the VRRP group resides or the IP address of an interface on a router in the VRRP group. In the latter case, the router is called the IP address owner.
- When a router is the IP address owner in a VRRP group, HP recommends you not to use the IP address of the interface (virtual IP address of the VRRP group) to establish a neighbor relationship with the adjacent router, that is, not to use the **network** command to enable OSPF on the interface. For more information about **network** command, see *Layer 3—IP Routing Command Reference*.
- When VRRP is operating in load balancing mode, the virtual IP address of a VRRP group cannot be the same as the IP address of any interface in the VRRP group. In other words, in load balancing mode, the VRRP group does not have an IP address owner.
- A VRRP group is removed after you remove all the virtual IP addresses configured for it. In addition, configurations on that VRRP group do not take effect any longer.
- Removal of the VRRP group on the IP address owner causes IP address collision. To solve the collision, modify the IP address of the interface on the IP address owner first and then remove the VRRP group from the interface.
- The virtual IP address of a VRRP group cannot be 0.0.0.0, 255.255.255.255, loopback addresses, non class A/B/C addresses or other illegal IP addresses such as 0.0.0.1.
- A VRRP group operates properly only when the configured virtual IP address and the interface IP address belong to the same segment and are legal host addresses. If the configured virtual IP address and the interface IP address do not belong to the same network segment, or the configured IP address is the network address or network broadcast address of the network segment to which the interface IP address belongs, the state of the VRRP group is always **initialize** though you can perform the configuration successfully. In this case, VRRP does not take effect.

## Configuration prerequisites

Before creating a VRRP group and configuring a virtual IP address on an interface, configure an IP address for the interface and make sure that it is in the same network segment as the virtual IP address to be configured.

## Configuration procedure

To create a VRRP group and configure a virtual IP address:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Enter the specified interface view.	<b>interface</b> <i>interface-type</i> <i>interface-number</i>	N/A
3. Create a VRRP group and configure a virtual IP address for the VRRP group.	<b>vrrp vrid</b> <i>virtual-router-id</i> <b>virtual-ip</b> <i>virtual-address</i>	VRRP group is not created by default.

# Configuring router priority, preemptive mode and tracking function

## Configuration guidelines

- The running priority of an IP address owner is always 255 and you do not need to configure it. An IP address owner always operates in preemptive mode.
- If you configure an interface to be tracked or a track entry to be monitored on a router that is the IP address owner in a VRRP group, the configuration does not take effect. If the router is not the IP address owner in the VRRP group later, the configuration takes effect.
- If the state of a tracked interface changes from down or removed to up, the priority of the router where the interface resides is automatically restored.
- If the state of a track entry changes from negative or invalid to positive, the priority of the router where the track entry is configured is automatically restored.

## Configuration prerequisites

Before you configure router priority, preemptive mode and tracking function, create a VRRP group on an interface and configure a virtual IP address for it.

## Configuration procedure

By configuring router priority, preemptive mode, interface tracking, or a track entry, you can determine which router in the VRRP group serves as the master.

To configure router priority, preemptive mode and the tracking function:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Enter interface view.	<b>interface</b> <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure router priority in the VRRP group.	<b>vrrp vrid</b> <i>virtual-router-id</i> <b>priority</b> <i>priority-value</i>	Optional. 100 by default.
4. Configure the router in the VRRP group to operate in preemptive mode and configure preemption delay.	<b>vrrp vrid</b> <i>virtual-router-id</i> <b>preempt-mode</b> [ <b>timer delay</b> <i>delay-value</i> ]	Optional. The router in the VRRP group operates in preemptive mode and the preemption delay is 0 seconds by default.
5. Configure the interface to be tracked.	<b>vrrp vrid</b> <i>virtual-router-id</i> <b>track</b> <b>interface</b> <i>interface-type</i> <i>interface-number</i> [ <b>reduced</b> <i>priority-reduced</i> ]	Optional. No interface is being tracked by default.
6. Configure VRRP to track a specified track entry.	<b>vrrp vrid</b> <i>virtual-router-id</i> <b>track</b> <i>track-entry-number</i> [ <b>reduced</b> <i>priority-reduced</i>   <b>switchover</b> ]	Optional. Not configured by default.

# Configuring VF tracking

## Configuration guidelines

- You can configure the VF tracking function when VRRP is operating in either standard protocol mode or load balancing mode. However, the VF tracking function is effective only when VRRP is operating in load balancing mode.
- By default, the weight of a VF is 255, and its lower limit of failure is 10.
- If the weight of a VF owner is higher than or equal to the lower limit of failure, the priority of the VF owner is always 255 and does not change with the weight value. Therefore, in case of an uplink failure, another VF takes over the VF owner's work and becomes the AVF only when the weight of the VF owner decreases by a properly specified value and becomes lower than the lower limit of failure. In other words, the weight of the VF owner decreases by more than 245.

## Configuration prerequisites

Before you configure the VF tracking function, create a VRRP group and configure a virtual IP address for it.

## Configuration procedure

VRRP operates in load balancing mode. Assume that you have configured the VF tracking function to monitor the track entry and specified the value by which the weight decreases. When the status of the track entry becomes negative, the weight values of all VFs on the router decrease by the specified value. When the status of the track entry becomes positive or invalid, the weight values of all VFs on the router restore their original values.

If you configure the VF tracking function on an LVF to monitor its corresponding AVF on a specified router, the LVF can take over the AVF immediately when the status of the track entry becomes negative, to ensure uninterrupted network communications.

To configure VF tracking:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Enter the specified interface view.	<b>interface</b> <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure VF tracking.	Configure the VF tracking function to monitor a specified track entry and specify the value by which the weight decreases: <b>vrrp vrid</b> <i>virtual-router-id</i> <b>weight</b> <b>track</b> <i>track-entry-number</i> [ <b>reduced</b> <i>weight-reduced</i> ] Configure the VF tracking function to monitor an AVF on a specified router: <b>vrrp vrid</b> <i>virtual-router-id</i> <b>track</b> <i>track-entry-number</i> <b>forwarder-switchover member-ip</b> <i>ip-address</i>	Use either approach. The VF tracking function is not configured by default.

# Configuring VRRP packet attributes

## Configuration guidelines

- You might configure different authentication modes and authentication keys for the VRRP groups on an interface. However, the members of the same VRRP group must use the same authentication mode and authentication key.
- Excessive traffic might cause a backup to trigger a change of its status because the backup does not receive any VRRP advertisements for a specified period of time. To solve this problem, prolong the time interval to send VRRP advertisements.
- Configuring different intervals for sending VRRP advertisements on the routers in a VRRP group might cause a backup to trigger a change of its status because the backup does not receive any VRRP advertisements for a specified period of time. To solve this problem, configure the same interval for sending VRRP advertisements on each router in the VRRP group.

## Configuration prerequisites

Before you configure the relevant attributes of VRRP packets, create a VRRP group and configure a virtual IP address for it.

## Configuration procedure

To configure VRRP packet attributes:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Set the Differentiated Services Code Point (DSCP) value for VRRP packets.	<b>vrrp dscp</b> <i>dscp-value</i>	Optional. 48 by default.
3. Enter the specified interface view.	<b>interface</b> <i>interface-type</i> <i>interface-number</i>	N/A
4. Configure the authentication mode and authentication key when the VRRP groups send and receive VRRP packets.	<b>vrrp vrid</b> <i>virtual-router-id</i> <b>authentication-mode</b> { <b>md5</b>   <b>simple</b> } [ <b>cipher</b> ] <i>key</i>	Optional. Authentication is not performed by default.
5. Configure the time interval for the master in the VRRP group to send VRRP advertisements.	<b>vrrp vrid</b> <i>virtual-router-id</i> <b>timer advertise</b> <i>adver-interval</i>	Optional. 1 second by default.
6. Disable TTL check on VRRP packets.	<b>vrrp un-check ttl</b>	Optional. Enabled by default. You do not need to create a VRRP group before executing this command.

## Enabling the trap function for VRRP

When the trap function is enabled for VRRP, VRRP generates traps with severity level **errors** to report its key events. The traps are sent to the information center of the device, where you can configure whether to output the trap information and the output destination. For how to configure the information center, see *Network Management and Monitoring Configuration Guide*.

To enable the trap function for VRRP:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Enable the trap function for VRRP.	<b>snmp-agent trap enable vrrp</b> [ <b>authfailure</b>   <b>newmaster</b> ]	Optional. Enabled by default.

For more information about the **snmp-agent trap enable vrrp** command, see the **snmp-agent trap enable** command in *Network Management and Monitoring Command Reference*.

## Displaying and maintaining VRRP for IPv4

Task	Command	Remarks
Display VRRP group status.	<b>display vrrp</b> [ <b>verbose</b> ] [ <b>interface</b> <i>interface-type interface-number</i> [ <b>vrid</b> <i>virtual-router-id</i> ] ] [ { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]	Available in any view
Display VRRP group statistics.	<b>display vrrp statistics</b> [ <b>interface</b> <i>interface-type interface-number</i> [ <b>vrid</b> <i>virtual-router-id</i> ] ] [ { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]	Available in any view
Clear VRRP group statistics.	<b>reset vrrp statistics</b> [ <b>interface</b> <i>interface-type interface-number</i> [ <b>vrid</b> <i>virtual-router-id</i> ] ]	Available in user view

## Configuring VRRP for IPv6

### VRRP for IPv6 configuration task list

Task	Remarks
Configuring a VRRP operation mode	Optional
Specifying the type of MAC addresses mapped to virtual IPv6 addresses	Optional When VRRP is operating in load balancing mode, this configuration is not effective.
Creating a VRRP group and configuring a virtual IPv6 address	Required
Configuring router priority, preemptive mode and tracking function	Optional
Configuring VF tracking	Optional The VF tracking function is effective only when VRRP is operating in load balancing mode.
Configuring VRRP packet attributes	Optional

## Specifying the type of MAC addresses mapped to virtual IPv6 addresses

After you specify the type of MAC addresses mapped to the virtual IPv6 address of VRRP groups and create a VRRP group, the master in the VRRP group uses the specified type of MAC address as the source MAC address for sending packets and uses the specified type of MAC address to answer ND requests from hosts so that the hosts in the internal network can learn the mapping between the IPv6 address and the MAC address.

The following types of MAC addresses are available to be mapped to the virtual IPv6 address of a VRRP group:

- **Virtual MAC to virtual IP mapping**—By default, a virtual MAC address is automatically created for a VRRP group when the VRRP group is created, and the virtual IPv6 address of the VRRP group is mapped to the virtual MAC address. When such a mapping is adopted, the hosts in the internal network do not need to update the mapping between the IPv6 address and the MAC address when the master changes.
- **Real MAC to virtual IP mapping**—In case that an IP address owner exists in a VRRP group, if the virtual IPv6 address is mapped to the virtual MAC address, two MAC addresses are mapped to one IPv6 address. To avoid such a problem, map the virtual IPv6 address of the VRRP group to the real MAC address of an interface to forward the packets from a host to the IP address owner.

When VRRP is operating in load balancing mode, a virtual IPv6 address is always mapped to a virtual MAC address regardless of which type of MAC addresses are specified to be mapped to virtual IPv6 addresses.

Specify the type of the MAC addresses mapped to the virtual IPv6 addresses before creating a VRRP group. Otherwise, you cannot change the type of the MAC addresses mapped to virtual IPv6 addresses.

To specify the type of MAC addresses mapped to virtual IPv6 addresses:

Step	Command	Remarks
1. Enter system view.	<code>system-view</code>	N/A
2. Specify the type of MAC addresses mapped to virtual IPv6 addresses.	<code>vrrp ipv6 method { real-mac   virtual-mac }</code>	Optional. Virtual MAC address by default.

## Creating a VRRP group and configuring a virtual IPv6 address

When creating a VRRP group, configure a virtual IPv6 address for the VRRP group. You can configure multiple virtual IPv6 addresses for a VRRP group.

A VRRP group is automatically created when you specify the first virtual IPv6 address for the VRRP group. If you specify another virtual IPv6 address for the VRRP group later, the virtual IPv6 address is added to the virtual IPv6 address list of the VRRP group.

Do not create VRRP groups on the VLAN interface of a super VLAN. Otherwise, network performance might be affected.

### Configuration guidelines

- When a router is the IP address owner in a VRRP group, HP recommends you not to use the IPv6 address of the interface (virtual IPv6 address of the VRRP group) to establish an OSPFv3 neighbor relationship with the adjacent router, that is, not to use the **ospfv3 area** command to enable

OSPFv3 on the interface. For more information about **ospfv3 area** command, see *Layer 3—IP Routing Command Reference*.

- When VRRP is operating in load balancing mode, the virtual IPv6 address of a VRRP group cannot be the same as the IPv6 address of any interface in the VRRP group. In other words, a VRRP group does not have an IP address owner in load balancing mode.
- A VRRP group is removed after you remove all the virtual IPv6 addresses in it. In addition, configurations on that VRRP group do not take effect any longer.
- Removal of the VRRP group on the IP address owner causes IP address collision. To resolve the collision, change the IPv6 address of the interface on the IP address owner first and then remove the VRRP group from the interface.

## Configuration prerequisites

Before creating a VRRP group and configuring a virtual IPv6 address on an interface, configure an IPv6 address for the interface and make sure that it is in the same network segment as the virtual IPv6 address to be configured.

## Configuration procedure

To create a VRRP group and configure its virtual IPv6 address:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Enter the specified interface view.	<b>interface</b> <i>interface-type</i> <i>interface-number</i>	N/A
3. Create a VRRP group and configure its virtual IPv6 address, which is a link local address.	<b>vrrp ipv6 vrid</b> <i>virtual-router-id</i> <b>virtual-ip</b> <i>virtual-address</i> <b>link-local</b>	No VRRP group is created by default. The first virtual IPv6 address of the VRRP group must be a link local address. Only one link local address is allowed in a VRRP group, and must be removed the last.
4. Configure the VRRP group with a virtual IPv6 address, which is a global unicast address.	<b>vrrp ipv6 vrid</b> <i>virtual-router-id</i> <b>virtual-ip</b> <i>virtual-address</i>	Optional. By default, no global unicast address is configured as the virtual IPv6 address of a VRRP group.

# Configuring router priority, preemptive mode and tracking function

## Configuration guidelines

- The running priority of an IP address owner is always 255 and you do not need to configure it. An IP address owner always operates in preemptive mode.
- Interface tracking is not configurable on an IP address owner.
- If you configure an interface to be tracked or a track entry to be monitored on a router that is the IP address owner in a VRRP group, the configuration does not take effect. If the router is not the IP address owner in the VRRP group later, the configuration takes effect.



- If the state of a tracked interface changes from down or removed to up, the priority of the router that owns the interface is automatically restored.
- If the state of a track entry changes from negative or invalid to positive, the priority of the router where the track entry is configured is automatically restored.

## Configuration prerequisites

Before you configure router priority, preemptive mode and tracking function, create a VRRP group and configure its virtual IPv6 address.

## Configuration procedure

By configuring router priority, preemptive mode, interface tracking, or a track entry, determine which router in the VRRP group serves as the master.

To configure router priority, preemptive mode and interface tracking:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Enter the specified interface view.	<b>interface</b> <i>interface-type</i> <i>interface-number</i>	N/A
3. Configure the priority of the router in the VRRP group.	<b>vrrp ipv6 vrid</b> <i>virtual-router-id</i> <b>priority</b> <i>priority-value</i>	Optional. 100 by default.
4. Configure the router in the VRRP group to operate in preemptive mode and configure preemption delay of the VRRP group.	<b>vrrp ipv6 vrid</b> <i>virtual-router-id</i> <b>preempt-mode</b> [ <b>timer delay</b> <i>delay-value</i> ]	Optional. The router in the VRRP group operates in preemptive mode and the preemption delay is zero seconds by default.
5. Configure the interface to be tracked.	<b>vrrp ipv6 vrid</b> <i>virtual-router-id</i> <b>track</b> <b>interface</b> <i>interface-type</i> <i>interface-number</i> [ <b>reduced</b> <i>priority-reduced</i> ]	Optional. No interface is being tracked by default.
6. Configure VRRP to track a specified track entry.	<b>vrrp ipv6 vrid</b> <i>virtual-router-id</i> <b>track</b> <i>track-entry-number</i> [ <b>reduced</b> <i>priority-reduced</i>   <b>switchover</b> ]	Optional. Not configured by default.

## Configuring VF tracking

### Configuration guidelines

- You can configure the VF tracking function when VRRP is operating in either standard protocol mode or load balancing mode. However, the VF tracking function is effective only when VRRP is operating in load balancing mode.
- By default, the weight of a VF is 255, and its lower limit of failure is 10.
- If the weight of a VF owner is higher than or equal to the lower limit of failure, the priority of the VF owner is always 255 and does not change with the weight value. Therefore, if an uplink fails, another VF takes over the VF owner's work and becomes the AVF only when the weight of the VF owner decreases by a properly specified value and becomes lower than the lower limit of failure. In other words, the weight of the VF owner decreases by more than 245.

## Configuration prerequisites

Before you configure the VF tracking function, create a VRRP group and configure a virtual IPv6 address for it.

## Configuration procedure

VRRP operates in load balancing mode. Assume that you have configured the VF tracking function to monitor a track entry and specified the value by which the weight decreases. When the status of the track entry becomes negative, the weight values of all VFs on the router decrease by the specified value. When the status of the track entry becomes positive or invalid, the weight values of all VFs on the router restore their original values.

If you configure the VF tracking function on an LVF to monitor its corresponding AVF on a specified router, the LVF can take over the AVF immediately when the status of the track entry becomes negative, to ensure uninterrupted network communications.

To configure VF tracking:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Enter the specified interface view.	<b>interface</b> <i>interface-type interface-number</i>	N/A
3. Configure VF tracking.	<p>Configure the VF tracking function to monitor a specified track entry and specify the value by which the weight decreases: <b>vrp ipv6 vrid</b> <i>virtual-router-id</i> <b>weight track</b> <i>track-entry-number</i> [ <b>reduced</b> <i>weight-reduced</i> ]</p> <p>Configure the VF tracking function to monitor an AVF on a specified router: <b>vrp ipv6 vrid</b> <i>virtual-router-id</i> <b>track</b> <i>track-entry-number</i> <b>forwarder-switchover</b> <b>member-ip</b> <i>ipv6-address</i></p>	<p>Use either approach. The VF tracking function is not configured by default.</p>

## Configuring VRRP packet attributes

### Configuration guidelines

- You might configure different authentication modes and authentication keys for the VRRP groups on an interface. However, the members of the same VRRP group must use the same authentication mode and authentication key.
- Excessive traffic might cause a backup to trigger a change of its status because the backup does not receive any VRRP advertisements for a specified period of time. To solve this problem, prolong the time interval to send VRRP advertisements.
- Configuring different intervals for sending VRRP advertisements on the routers in a VRRP group might cause a backup to trigger a change of its status because the backup does not receive any VRRP advertisements for a specified period of time. To solve this problem, configure the same interval for sending VRRP advertisements on each router in the VRRP group.

### Configuration prerequisites

Before you configure the relevant attributes of VRRP packets, create a VRRP group and configure a virtual IPv6 address.

## Configuration procedure

To configure VRRP packet attributes:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Set the DSCP value for VRRP packets.	<b>vrrp ipv6 dscp</b> <i>dscp-value</i>	Optional. 56 by default.
3. Enter the specified interface view.	<b>interface</b> <i>interface-type</i> <i>interface-number</i>	N/A
4. Configure the authentication mode and authentication key when the VRRP groups send or receive VRRP packets.	<b>vrrp ipv6 vrid</b> <i>virtual-router-id</i> <b>authentication-mode</b> <b>simple</b> [ <b>cipher</b> ] <i>key</i>	Optional. Authentication is not performed by default.
5. Configure the time interval for the master in the VRRP group to send VRRP advertisement.	<b>vrrp ipv6 vrid</b> <i>virtual-router-id</i> <b>timer</b> <b>advertise</b> <i>adver-interval</i>	Optional. 100 centiseconds by default.

## Displaying and maintaining VRRP for IPv6

Task	Command	Remarks
Display VRRP group status.	<b>display vrrp ipv6</b> [ <b>verbose</b> ] [ <b>interface</b> <i>interface-type interface-number</i> [ <b>vrid</b> <i>virtual-router-id</i> ] ] [ { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]	Available in any view
Display VRRP group statistics.	<b>display vrrp ipv6 statistics</b> [ <b>interface</b> <i>interface-type interface-number</i> [ <b>vrid</b> <i>virtual-router-id</i> ] ] [ { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]	Available in any view
Clear VRRP group statistics.	<b>reset vrrp ipv6 statistics</b> [ <b>interface</b> <i>interface-type</i> <i>interface-number</i> [ <b>vrid</b> <i>virtual-router-id</i> ] ]	Available in user view

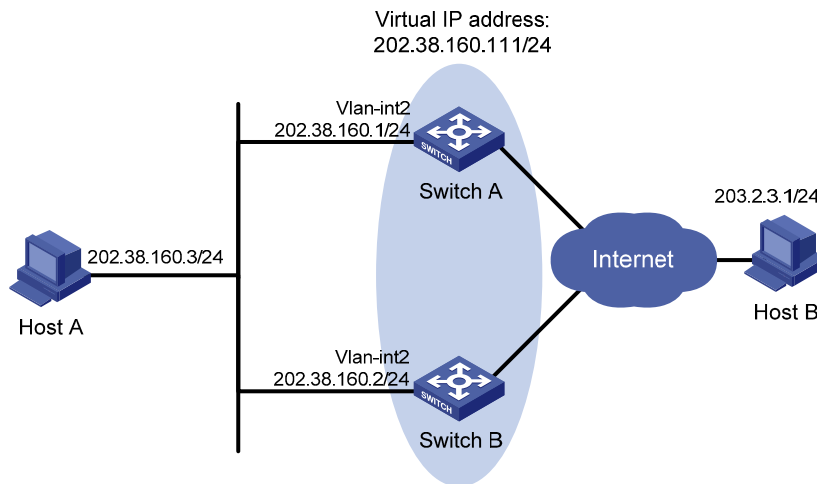
## IPv4-based VRRP configuration examples

### Single VRRP group configuration example

#### Network requirements

- Host A wants to access Host B on the Internet, using 202.38.160.111/24 as its default gateway.
- Switch A and Switch B belong to VRRP group 1 with the virtual IP address of 202.38.160.111/24.
- When Switch A operates normally, packets sent from Host A to Host B are forwarded by Switch A; when Switch A fails, packets sent from Host A to Host B are forwarded by Switch B.

**Figure 40 Network diagram**



## Configuration procedure

### 1. Configure Switch A:

# Configure VLAN 2.

```
<SwitchA> system-view
[SwitchA] vlan 2
[SwitchA-vlan2] port gigabitethernet 1/0/5
[SwitchA-vlan2] quit
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ip address 202.38.160.1 255.255.255.0
```

# Create VRRP group 1 and set its virtual IP address to 202.38.160.111.

```
[SwitchA-Vlan-interface2] vrrp vrid 1 virtual-ip 202.38.160.111
```

# Set the priority of Switch A in VRRP group 1 to 110, which is higher than that of Switch B (100), so that Switch A can become the master.

```
[SwitchA-Vlan-interface2] vrrp vrid 1 priority 110
```

# Configure Switch A to operate in preemptive mode so that it can become the master whenever it works normally, and configure the preempt delay as five seconds to avoid frequent status switchover.

```
[SwitchA-Vlan-interface2] vrrp vrid 1 preempt-mode timer delay 5
```

### 2. Configure Switch B:

# Configure VLAN 2.

```
<SwitchB> system-view
[SwitchB] vlan 2
[SwitchB-Vlan2] port gigabitethernet 1/0/5
[SwitchB-vlan2] quit
[SwitchB] interface vlan-interface 2
[SwitchB-Vlan-interface2] ip address 202.38.160.2 255.255.255.0
```

# Create VRRP group 1 and set its virtual IP address to 202.38.160.111.

```
[SwitchB-Vlan-interface2] vrrp vrid 1 virtual-ip 202.38.160.111
```

# Set Switch B to operate in preemptive mode. The preempt delay is five seconds.

```
[SwitchB-Vlan-interface2] vrrp vrid 1 preempt-mode timer delay 5
```

### 3. Verify the configuration:

After the configuration, Host B can be pinged successfully on Host A. To verify your configuration, use the **display vrrp verbose** command.

# Display the detailed information about VRRP group 1 on Switch A.

```
[SwitchA-Vlan-interface2] display vrrp verbose
IPv4 Standby Information:
 Run Mode : Standard
 Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 1
 Admin Status : Up State : Master
 Config Pri : 110 Running Pri : 110
 Preempt Mode : Yes Delay Time : 5
 Auth Type : None
 Virtual IP : 202.38.160.111
 Virtual MAC : 0000-5e00-0101
 Master IP : 202.38.160.1
```

# Display the detailed information about VRRP group 1 on Switch B.

```
[SwitchB-Vlan-interface2] display vrrp verbose
IPv4 Standby Information:
 Run Mode : Standard
 Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 1
 Admin Status : Up State : Backup
 Config Pri : 100 Running Pri : 100
 Preempt Mode : Yes Delay Time : 5
 Become Master : 4200ms left
 Auth Type : None
 Virtual IP : 202.38.160.111
 Master IP : 202.38.160.1
```

The output shows that in VRRP group 1 Switch A is the master, Switch B is the backup and packets sent from Host A to Host B are forwarded by Switch A.

When Switch A fails, you can still ping through Host B on Host A. Use the **display vrrp verbose** command to view the detailed information about the VRRP group on Switch B.

# When Switch A fails, the detailed information about VRRP group 1 on Switch B is displayed.

```
[SwitchB-Vlan-interface2] display vrrp verbose
IPv4 Standby Information:
 Run Mode : Standard
 Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 1
 Admin Status : Up State : Master
 Config Pri : 100 Running Pri : 100
```

```
Preempt Mode : Yes Delay Time : 5
Auth Type : None
Virtual IP : 202.38.160.111
Virtual MAC : 0000-5e00-0101
Master IP : 202.38.160.2
```

The output shows that when Switch A fails, Switch B becomes the master, and packets sent from Host A to Host B are forwarded by Switch B.

# After Switch A resumes normal operation, use the **display vrrp verbose** command to display the detailed information about VRRP group 1 on Switch A.

```
[SwitchA-Vlan-interface2] display vrrp verbose
IPv4 Standby Information:
 Run Mode : Standard
 Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 1
 Admin Status : Up State : Master
 Config Pri : 110 Running Pri : 110
 Preempt Mode : Yes Delay Time : 5
 Auth Type : None
 Virtual IP : 202.38.160.111
 Virtual MAC : 0000-5e00-0101
 Master IP : 202.38.160.1
```

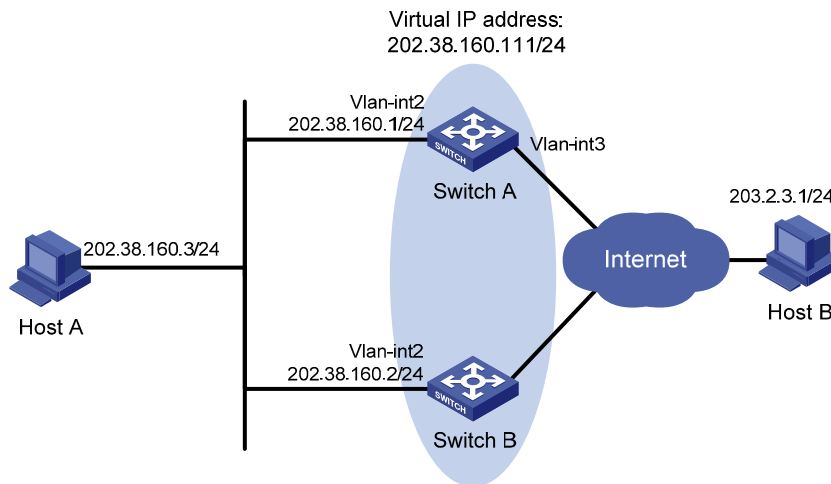
The output shows that after Switch A resumes normal operation, it becomes the master, and packets sent from host A to host B are forwarded by Switch A.

## VRRP interface tracking configuration example

### Network requirements

- Host A wants to access Host B on the Internet, using 202.38.160.111/24 as its default gateway.
- Switch A and Switch B belong to VRRP group 1 with the virtual IP address of 202.38.160.111/24.
- If Switch A operates normally, packets sent from Host A to Host B are forwarded by Switch A. If VLAN-interface 3 through which Switch A connects to the Internet is not available, packets sent from Host A to Host B are forwarded by Switch B.
- To prevent attacks to the VRRP group from illegal users by using spoofed packets, configure the authentication mode as plain text to authenticate the VRRP packets in VRRP group 1, and specify the authentication key as **hello**.

Figure 41 Network diagram



## Configuration procedure

### 1. Configure Switch A:

# Configure VLAN 2.

```
<SwitchA> system-view
[SwitchA] vlan 2
[SwitchA-vlan2] port gigabitethernet 1/0/5
[SwitchA-vlan2] quit
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ip address 202.38.160.1 255.255.255.0
```

# Create a VRRP group 1 and set its virtual IP address to 202.38.160.111.

```
[SwitchA-Vlan-interface2] vrrp vrid 1 virtual-ip 202.38.160.111
```

# Configure the priority of Switch A in the VRRP group to 110, which is higher than that of Switch B (100), so that Switch A can become the master.

```
[SwitchA-Vlan-interface2] vrrp vrid 1 priority 110
```

# Configure the authentication mode of the VRRP group as **simple** and authentication key as **hello**.

```
[SwitchA-Vlan-interface2] vrrp vrid 1 authentication-mode simple hello
```

# Set the interval for Master to send VRRP advertisement to four seconds.

```
[SwitchA-Vlan-interface2] vrrp vrid 1 timer advertise 4
```

# Configure Switch A to operate in preemptive mode, so that it can become the master whenever it works normally; configure the preemption delay as five seconds to avoid frequent status switchover.

```
[SwitchA-Vlan-interface2] vrrp vrid 1 preempt-mode timer delay 5
```

# Set VLAN interface 3 on Switch A to be tracked, and configure the amount by which the priority value decreases to be more than 10 (30 in this example), so that when VLAN-interface 3 fails, the priority of Switch A in VRRP group 1 decreases to a value lower than 100 and thus Switch B can become the master.

```
[SwitchA-Vlan-interface2] vrrp vrid 1 track interface vlan-interface 3 reduced 30
```

### 2. Configure Switch B:

# Configure VLAN 2.

```
<SwitchB> system-view
[SwitchB] vlan 2
```

```

[SwitchB-vlan2] port gigabitethernet 1/0/5
[SwitchB-vlan2] quit
[SwitchB] interface vlan-interface 2
[SwitchB-Vlan-interface2] ip address 202.38.160.2 255.255.255.0
Create a VRRP group 1 and set its virtual IP address to 202.38.160.111.
[SwitchB-Vlan-interface2] vrrp vrid 1 virtual-ip 202.38.160.111
Configure the authentication mode of the VRRP group as simple and authentication key as hello.
[SwitchB-Vlan-interface2] vrrp vrid 1 authentication-mode simple hello
Set the interval for master to send VRRP advertisement to four seconds.
[SwitchB-Vlan-interface2] vrrp vrid 1 timer advertise 4
Configure Switch B to operate in preemptive mode, so that Switch B can become the master after
the priority of Switch A decreases to a value lower than 100. Configure the preemption delay as
five seconds to avoid frequent status switchover.
[SwitchB-Vlan-interface2] vrrp vrid 1 preempt-mode timer delay 5

```

### 3. Verify the configuration:

After the configuration, Host B can be pinged successfully on Host A. To verify your configuration, use the **display vrrp verbose** command.

# Display the detailed information about VRRP group 1 on Switch A.

```

[SwitchA-Vlan-interface2] display vrrp verbose
IPv4 Standby Information:
 Run Mode : Standard
 Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 4
 Admin Status : Up State : Master
 Config Pri : 110 Running Pri : 110
 Preempt Mode : Yes Delay Time : 5
 Auth Type : Simple Key : *****
 Virtual IP : 202.38.160.111
 Virtual MAC : 0000-5e00-0101
 Master IP : 202.38.160.1
VRRP Track Information:
 Track Interface: Vlan3 State : Up Pri Reduced : 30

```

# Display the detailed information about VRRP group 1 on Switch B.

```

[SwitchB-Vlan-interface2] display vrrp verbose
IPv4 Standby Information:
 Run Mode : Standard
 Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 4
 Admin Status : Up State : Backup
 Config Pri : 100 Running Pri : 100
 Preempt Mode : Yes Delay Time : 5
 Become Master : 2200ms left
 Auth Type : Simple Key : *****

```



```
Virtual IP : 202.38.160.111
Master IP : 202.38.160.1
```

The output shows that in VRRP group 1 Switch A is the master, Switch B is the backup and packets sent from Host A to Host B are forwarded by Switch A.

If interface VLAN-interface 3 through which Switch A connects to the Internet is not available, you can still ping Host B successfully on Host A. To view the detailed information about the VRRP group, use the **display vrrp verbose** command.

# If VLAN-interface 3 on Switch A is not available, the detailed information about VRRP group 1 on Switch A is displayed.

```
[SwitchA-Vlan-interface2] display vrrp verbose
IPv4 Standby Information:
 Run Mode : Standard
 Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 4
 Admin Status : Up State : Backup
 Config Pri : 110 Running Pri : 80
 Preempt Mode : Yes Delay Time : 5
 Become Master : 2200ms left
 Auth Type : Simple Key : *****
 Virtual IP : 202.38.160.111
 Master IP : 202.38.160.2
VRRP Track Information:
 Track Interface: Vlan3 State : Down Pri Reduced : 30
```

# When VLAN-interface 3 on Switch A is not available, the detailed information about VRRP group 1 on Switch B is displayed.

```
[SwitchB-Vlan-interface2] display vrrp verbose
IPv4 Standby Information:
 Run Mode : Standard
 Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 4
 Admin Status : Up State : Master
 Config Pri : 100 Running Pri : 100
 Preempt Mode : Yes Delay Time : 5
 Auth Type : Simple Key : *****
 Virtual IP : 202.38.160.111
 Virtual MAC : 0000-5e00-0101
 Master IP : 202.38.160.2
```

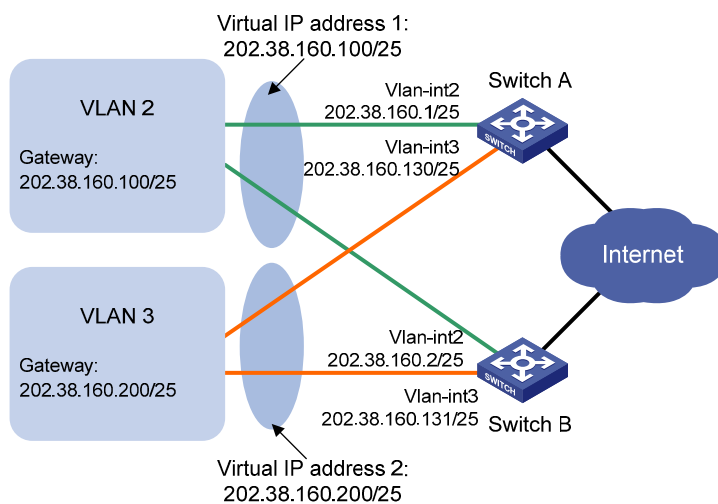
The output shows that when VLAN-interface 3 on Switch A is not available, the priority of Switch A is reduced to 80 and it becomes the backup. Switch B becomes the master and packets sent from Host A to Host B are forwarded by Switch B.

# VRRP with multiple VLANs configuration example

## Network requirements

- Hosts in VLAN 2 use 202.38.160.100/25 as their default gateway and hosts in VLAN 3 use 202.38.160.200/25 as their default gateway.
- Switch A and Switch B belong to both VRRP group 1 and VRRP group 2. The virtual IP address of VRRP group 1 is 202.38.160.100/25, and that of VRRP group 2 is 202.38.160.200/25.
- In VRRP group 1, Switch A has a higher priority than Switch B. In VRRP group 2, Switch B has a higher priority than Switch A. In this case, hosts in VLAN 2 and VLAN 3 can communicate with external networks through Switch A and Switch B, respectively, and when Switch A or Switch B fails, the hosts can use the other switch to communicate with external networks to avoid communication interruption.

Figure 42 Network diagram



## Configuration procedure

### 1. Configure Switch A:

# Configure VLAN 2.

```
<SwitchA> system-view
[SwitchA] vlan 2
[SwitchA-vlan2] port gigabitethernet 1/0/5
[SwitchA-vlan2] quit
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ip address 202.38.160.1 255.255.255.128
```

# Create a VRRP group 1 and set its virtual IP address to 202.38.160.100.

```
[SwitchA-Vlan-interface2] vrrp vrid 1 virtual-ip 202.38.160.100
```

# Configure the priority of Switch A in VRRP group 1 as 110, which is higher than that of Switch B (100), so that Switch A can become the master in VRRP group 1.

```
[SwitchA-Vlan-interface2] vrrp vrid 1 priority 110
```

```
[SwitchA-Vlan-interface2] quit
```

# Configure VLAN 3.

```
[SwitchA] vlan 3
[SwitchA-vlan3] port gigabitethernet 1/0/6
```

```
[SwitchA-vlan3] quit
[SwitchA] interface vlan-interface 3
[SwitchA-Vlan-interface3] ip address 202.38.160.130 255.255.255.128
Create a VRRP group 2 and set its virtual IP address to 202.38.160.200.
[SwitchA-Vlan-interface3] vrrp vrid 2 virtual-ip 202.38.160.200
```

## 2. Configure Switch B:

### # Configure VLAN 2.

```
<SwitchB> system-view
[SwitchB] vlan 2
[SwitchB-vlan2] port gigabitethernet 1/0/5
[SwitchB-vlan2] quit
[SwitchB] interface vlan-interface 2
[SwitchB-Vlan-interface2] ip address 202.38.160.2 255.255.255.128
Create a VRRP group 1 and set its virtual IP address to 202.38.160.100.
[SwitchB-Vlan-interface2] vrrp vrid 1 virtual-ip 202.38.160.100
[SwitchB-Vlan-interface2] quit
```

### # Configure VLAN 3.

```
[SwitchB] vlan 3
[SwitchB-vlan3] port gigabitethernet 1/0/6
[SwitchB-vlan3] quit
[SwitchB] interface vlan-interface 3
[SwitchB-Vlan-interface3] ip address 202.38.160.131 255.255.255.128
Create a VRRP group 2 and set its virtual IP address to 202.38.160.200.
[SwitchB-Vlan-interface3] vrrp vrid 2 virtual-ip 202.38.160.200
Configure the priority of Switch B in VRRP group 2 to 110, which is higher than that of Switch A (100), so that Switch B can become the master in VRRP group 2.
[SwitchB-Vlan-interface3] vrrp vrid 2 priority 110
```

## 3. Verify the configuration:

To verify your configuration, use the **display vrrp verbose** command.

# Display the detailed information about the VRRP group on Switch A.

```
[SwitchA-Vlan-interface3] display vrrp verbose
IPv4 Standby Information:
 Run Mode : Standard
 Run Method : Virtual MAC
Total number of virtual routers : 2
Interface Vlan-interface2
 VRID : 1 Adver Timer : 1
 Admin Status : Up State : Master
 Config Pri : 110 Running Pri : 110
 Preempt Mode : Yes Delay Time : 0
 Auth Type : None
 Virtual IP : 202.38.160.100
 Virtual MAC : 0000-5e00-0101
 Master IP : 202.38.160.1
Interface Vlan-interface3
 VRID : 2 Adver Timer : 1
```

```

Admin Status : Up State : Backup
Config Pri : 100 Running Pri : 100
Preempt Mode : Yes Delay Time : 0
Become Master : 2200ms left
Auth Type : None
Virtual IP : 202.38.160.200
Master IP : 202.38.160.131

```

# Display the detailed information about the VRRP group on Switch B.

```

[SwitchB-Vlan-interface3] display vrrp verbose
IPv4 Standby Information:
 Run Mode : Standard
 Run Method : Virtual MAC
Total number of virtual routers : 2
Interface Vlan-interface2
 VRID : 1 Adver Timer : 1
 Admin Status : Up State : Backup
 Config Pri : 100 Running Pri : 100
 Preempt Mode : Yes Delay Time : 0
 Become Master : 2200ms left
 Auth Type : None
 Virtual IP : 202.38.160.100
 Master IP : 202.38.160.1
Interface Vlan-interface3
 VRID : 2 Adver Timer : 1
 Admin Status : Up State : Master
 Config Pri : 110 Running Pri : 110
 Preempt Mode : Yes Delay Time : 0
 Auth Type : None
 Virtual IP : 202.38.160.200
 Virtual MAC : 0000-5e00-0102
 Master IP : 202.38.160.131

```

The output shows that in VRRP group 1 Switch A is the master, Switch B is the backup and hosts with the default gateway of 202.38.160.100/25 accesses the Internet through Switch A; in VRRP group 2 Switch A is the backup, Switch B is the master and hosts with the default gateway of 202.38.160.200/25 accesses the Internet through Switch B.

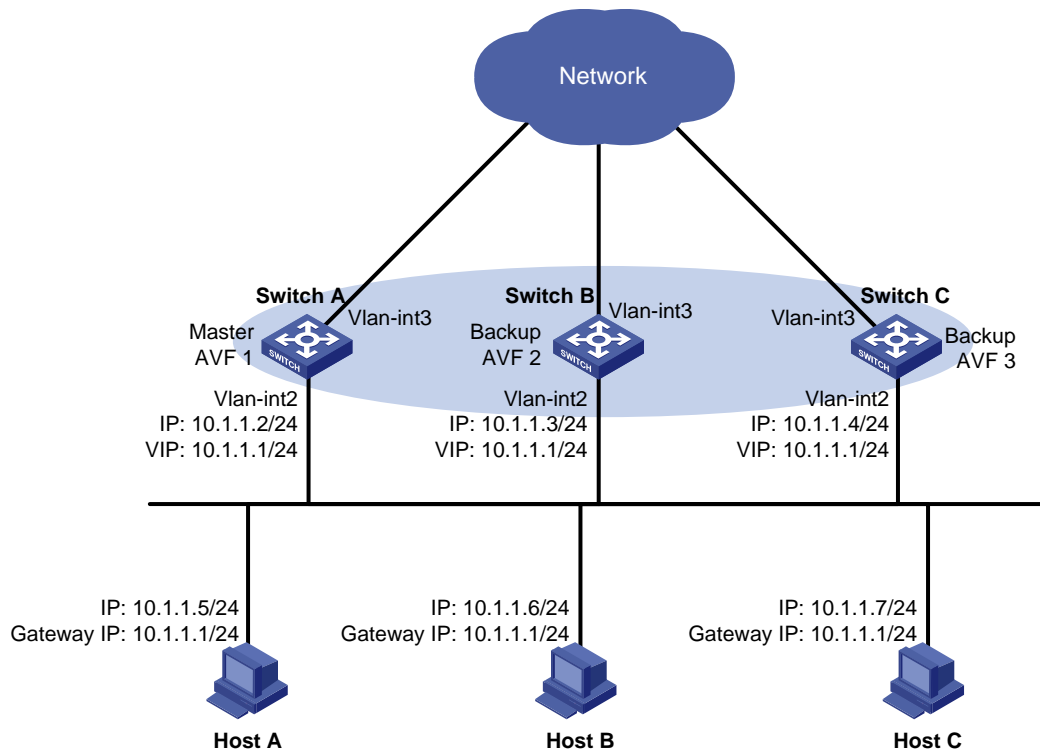
## VRRP load balancing mode configuration example

### Network requirements

- Switch A, Switch B, and Switch C belong to VRRP group 1 with the virtual IP address of 10.1.1.1/24.
- Hosts on network segment 10.1.1.0/24 use 10.1.1.1/24 as their default gateway. Use the VRRP group to make sure that when a gateway (Switch A, Switch B, or Switch C) fails, the hosts on the LAN can access external networks through another gateway.
- VRRP group 1 is operating in load balancing mode to make good use of network resources.
- Configure a track entry on Switch A, Switch B, and Switch C to monitor their own VLAN-interface 3. When the interface on Switch A, Switch B, or Switch C fails, the weight of the corresponding switch decreases so that another switch with a higher weight can take over.

- Configure track entries on Switch C to monitor Switch A and Switch B. When Switch A or Switch B fails, Switch C immediately takes over the AVF on Switch A or Switch B.

Figure 43 Network diagram



## Configuration procedure

### 1. Configure Switch A:

# Configure VLAN 2.

```
<SwitchA> system-view
[SwitchA] vlan 2
[SwitchA-vlan2] port gigabitethernet 1/0/5
[SwitchA-vlan2] quit
```

# Configure VRRP to operate in load balancing mode.

```
[SwitchA] vrrp mode load-balance
```

# Create VRRP group 1 and configure its virtual IP address as 10.1.1.1.

```
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ip address 10.1.1.2 24
[SwitchA-Vlan-interface2] vrrp vrid 1 virtual-ip 10.1.1.1
```

# Set the priority of Switch A in VRRP group 1 to 120, which is higher than that of Switch B (110) and that of Switch C (100), so that Switch A can become the master.

```
[SwitchA-Vlan-interface2] vrrp vrid 1 priority 120
```

# Configure Switch A to operate in preemptive mode, so that it can become the master whenever it works normally; configure the preemption delay as five seconds to avoid frequent status switchover.

```
[SwitchA-Vlan-interface2] vrrp vrid 1 preempt-mode timer delay 5
[SwitchA-Vlan-interface2] quit
```

# Create track entry 1 to associate with the physical status of VLAN-interface 3 on Switch A. When the track entry becomes negative, it means that the interface fails.

```
[SwitchA] track 1 interface vlan-interface 3
```

# Configure VF tracking to monitor track entry 1 and specify the value by which the weight decreases, making the weight of Switch A decrease by more than 245 (250 in this example) when track entry 1 turns to negative. In such a case, another router with a higher weight can take over.

```
[SwitchA] interface vlan-interface 2
```

```
[SwitchA-Vlan-interface2] vrrp vrid 1 weight track 1 reduced 250
```

## 2. Configure Switch B:

# Configure VLAN 2.

```
<SwitchB> system-view
```

```
[SwitchB] vlan 2
```

```
[SwitchB-vlan2] port gigabitethernet 1/0/5
```

```
[SwitchB-vlan2] quit
```

# Configure VRRP to operate in load balancing mode.

```
[SwitchB] vrrp mode load-balance
```

# Create VRRP group 1 and configure its virtual IP address as 10.1.1.1.

```
[SwitchB] interface vlan-interface 2
```

```
[SwitchB-Vlan-interface2] ip address 10.1.1.3 24
```

```
[SwitchB-Vlan-interface2] vrrp vrid 1 virtual-ip 10.1.1.1
```

# Set the priority of Switch B in VRRP group 1 to 110, which is higher than that of Switch C (100), so that Switch B can become the master when Switch A fails.

```
[SwitchB-Vlan-interface2] vrrp vrid 1 priority 110
```

# Set Switch B to operate in preemptive mode. The preemption delay is five seconds.

```
[SwitchB-Vlan-interface2] vrrp vrid 1 preempt-mode timer delay 5
```

```
[SwitchB-Vlan-interface2] quit
```

# Create track entry 1 to associate with the physical status of VLAN-interface 3 on Switch B. When the track entry becomes negative, it means that the interface fails.

```
[SwitchB] track 1 interface vlan-interface 3
```

# Configure VF tracking to monitor track entry 1 and specify the value by which the weight decreases, making the weight of Switch B decrease by more than 245 (250 in this example) when track entry 1 turns to negative. In such a case, another router with a higher weight can take over.

```
[SwitchB] interface vlan-interface 2
```

```
[SwitchB-Vlan-interface2] vrrp vrid 1 weight track 1 reduced 250
```

## 3. Configure Switch C:

# Configure VLAN 2.

```
<SwitchC> system-view
```

```
[SwitchC] vlan 2
```

```
[SwitchC-vlan2] port gigabitethernet 1/0/5
```

```
[SwitchC-vlan2] quit
```

# Configure VRRP to operate in load balancing mode.

```
[SwitchC] vrrp mode load-balance
```

# Create VRRP group 1 and configure its virtual IP address as 10.1.1.1.

```
[SwitchC] interface vlan-interface 2
```

```
[SwitchC-Vlan-interface2] ip address 10.1.1.4 24
```

```
[SwitchC-Vlan-interface2] vrrp vrid 1 virtual-ip 10.1.1.1
```

# Set Switch C to operate in preemptive mode. The preemption delay is five seconds.

```
[SwitchC-Vlan-interface2] vrrp vrid 1 preempt-mode timer delay 5
[SwitchC-Vlan-interface2] quit
```

# Create track entry 1 to associate with the physical status of VLAN-interface 3 on Switch C. When the track entry becomes negative, it means that the interface fails.

```
[SwitchC] track 1 interface vlan-interface 3
```

# Configure VF tracking to monitor track entry 1 and specify the value by which the weight decreases, making the weight of Switch C decrease by more than 245 (250 in this example) when track entry 1 turns to negative. In such a case, another router with a higher weight can take over.

```
[SwitchC] interface vlan-interface 2
[SwitchC-Vlan-interface2] vrrp vrid 1 weight track 1 reduced 250
[SwitchC-Vlan-interface2] quit
```

# Create track entries 2 and 3 to monitor Switch A and Switch B, respectively. If a track entry becomes negative, it indicates that the corresponding switch fails.

```
[SwitchC] bfd echo-source-ip 1.2.3.4
[SwitchC] track 2 bfd echo interface vlan-interface 2 remote ip 10.1.1.2 local ip 10.1.1.4
[SwitchC] track 3 bfd echo interface vlan-interface 2 remote ip 10.1.1.3 local ip 10.1.1.4
```

# Configure VF tracking to monitor track entry 2. When track entry 2 becomes negative, the LVF on Switch C whose corresponding AVF is on the switch with the IP address of 10.1.1.2 immediately becomes active. Switch C takes over the AVF on Switch A.

```
[SwitchC] interface vlan-interface 2
[SwitchC-Vlan-interface2] vrrp vrid 1 track 2 forwarder-switchover member-ip 10.1.1.2
```

# Configure VF tracking to monitor track entry 3. When track entry 2 becomes negative, the LVF on Switch C whose corresponding AVF is on the switch with the IP address of 10.1.1.3 immediately becomes active. Switch C takes over the AVF on Switch B.

```
[SwitchC-Vlan-interface2] vrrp vrid 1 track 3 forwarder-switchover member-ip 10.1.1.3
```

#### 4. Verify the configuration:

After the configuration, Host A can ping the external network. To verify your configuration, use the **display vrrp verbose** command.

# Display the detailed information about VRRP group 1 on Switch A.

```
[SwitchA-Vlan-interface2] display vrrp verbose
IPv4 Standby Information:
 Run Mode : Load Balance
 Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 1
 Admin Status : Up State : Master
 Config Pri : 120 Running Pri : 120
 Preempt Mode : Yes Delay Time : 5
 Auth Type : None
 Virtual IP : 10.1.1.1
 Member IP List : 10.1.1.2 (Local, Master)
 10.1.1.3 (Backup)
 10.1.1.4 (Backup)
```

Forwarder Information: 3 Forwarders 1 Active

Config Weight : 255  
Running Weight : 255

Forwarder 01

State : Active  
Virtual MAC : 000f-e2ff-0011 (Owner)  
Owner ID : 0000-5e01-1101  
Priority : 255  
Active : local

Forwarder 02

State : Listening  
Virtual MAC : 000f-e2ff-0012 (Learnt)  
Owner ID : 0000-5e01-1103  
Priority : 127  
Active : 10.1.1.3

Forwarder 03

State : Listening  
Virtual MAC : 000f-e2ff-0013 (Learnt)  
Owner ID : 0000-5e01-1105  
Priority : 127  
Active : 10.1.1.4

Forwarder Weight Track Information:

Track Object : 1 State : Positive Weight Reduced : 250

# Display the detailed information about VRRP group 1 on Switch B.

[SwitchB-Vlan-interface2] display vrrp verbose

IPv4 Standby Information:

Run Mode : Load Balance  
Run Method : Virtual MAC

Total number of virtual routers : 1

Interface Vlan-interface2

VRID : 1 Adver Timer : 1  
Admin Status : Up State : Backup  
Config Pri : 110 Running Pri : 110  
Preempt Mode : Yes Delay Time : 5  
Become Master : 4200ms left  
Auth Type : None  
Virtual IP : 10.1.1.1  
Member IP List : 10.1.1.3 (Local, Backup)  
10.1.1.2 (Master)  
10.1.1.4 (Backup)

Forwarder Information: 3 Forwarders 1 Active

Config Weight : 255  
Running Weight : 255

Forwarder 01

State : Listening  
Virtual MAC : 000f-e2ff-0011 (Learnt)  
Owner ID : 0000-5e01-1101  
Priority : 127



Active : 10.1.1.2

Forwarder 02

State : Active  
Virtual MAC : 000f-e2ff-0012 (Owner)  
Owner ID : 0000-5e01-1103  
Priority : 255  
Active : local

Forwarder 03

State : Listening  
Virtual MAC : 000f-e2ff-0013 (Learnt)  
Owner ID : 0000-5e01-1105  
Priority : 127  
Active : 10.1.1.4

Forwarder Weight Track Information:

Track Object : 1 State : Positive Weight Reduced : 250

# Display the detailed information about VRRP group 1 on Switch C.

[SwitchC-Vlan-interface2] display vrrp verbose

IPv4 Standby Information:

Run Mode : Load Balance  
Run Method : Virtual MAC

Total number of virtual routers : 1

Interface Vlan-interface2

VRID : 1 Adver Timer : 1  
Admin Status : Up State : Backup  
Config Pri : 100 Running Pri : 100  
Preempt Mode : Yes Delay Time : 5  
Become Master : 4200ms left  
Auth Type : None  
Virtual IP : 10.1.1.1  
Member IP List : 10.1.1.4 (Local, Backup)  
10.1.1.2 (Master)  
10.1.1.3 (Backup)

Forwarder Information: 3 Forwarders 1 Active

Config Weight : 255  
Running Weight : 255

Forwarder 01

State : Listening  
Virtual MAC : 000f-e2ff-0011 (Learnt)  
Owner ID : 0000-5e01-1101  
Priority : 127  
Active : 10.1.1.2

Forwarder 02

State : Listening  
Virtual MAC : 000f-e2ff-0012 (Learnt)  
Owner ID : 0000-5e01-1103  
Priority : 127  
Active : 10.1.1.3

Forwarder 03

```

State : Active
Virtual MAC : 000f-e2ff-0013 (Owner)
Owner ID : 0000-5e01-1105
Priority : 255
Active : local
Forwarder Weight Track Information:
Track Object : 1 State : Positive Weight Reduced : 250
Forwarder Switchover Track Information:
Track Object : 2 State : Positive
Member IP : 10.1.1.2
Track Object : 3 State : Positive
Member IP : 10.1.1.3

```

The output shows that in VRRP group 1, Switch A is the master and Switch B and Switch C are the backups. Each switch has one AVF and two LVFs that act as the backups.

# When VLAN-interface 3 on Switch A fails, use the **display vrrp verbose** command to display the detailed information about VRRP group 1 on Switch A.

```

[SwitchA-Vlan-interface2] display vrrp verbose
IPv4 Standby Information:
Run Mode : Load Balance
Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
VRID : 1 Adver Timer : 1
Admin Status : Up State : Master
Config Pri : 120 Running Pri : 120
Preempt Mode : Yes Delay Time : 5
Auth Type : None
Virtual IP : 10.1.1.1
Member IP List : 10.1.1.2 (Local, Master)
 10.1.1.3 (Backup)
 10.1.1.4 (Backup)
Forwarder Information: 3 Forwarders 0 Active
Config Weight : 255
Running Weight : 5
Forwarder 01
State : Initialize
Virtual MAC : 000f-e2ff-0011 (Owner)
Owner ID : 0000-5e01-1101
Priority : 0
Active : 10.1.1.4
Forwarder 02
State : Initialize
Virtual MAC : 000f-e2ff-0012 (Learnt)
Owner ID : 0000-5e01-1103
Priority : 0
Active : 10.1.1.3
Forwarder 03
State : Initialize

```

```
Virtual MAC : 000f-e2ff-0013 (Learnt)
Owner ID : 0000-5e01-1105
Priority : 0
Active : 10.1.1.4
```

Forwarder Weight Track Information:

```
Track Object : 1 State : Negative Weight Reduced : 250
```

# Use the **display vrrp verbose** command to display the detailed information about VRRP group 1 on Switch C.

```
[SwitchC-Vlan-interface2] display vrrp verbose
```

IPv4 Standby Information:

```
Run Mode : Load Balance
Run Method : Virtual MAC
```

Total number of virtual routers : 1

Interface Vlan-interface2

```
VRID : 1 Adver Timer : 1
Admin Status : Up State : Backup
Config Pri : 100 Running Pri : 100
Preempt Mode : Yes Delay Time : 5
Become Master : 4200ms left
Auth Type : None
Virtual IP : 10.1.1.1
Member IP List : 10.1.1.4 (Local, Backup)
 10.1.1.2 (Master)
 10.1.1.3 (Backup)
```

Forwarder Information: 3 Forwarders 2 Active

```
Config Weight : 255
Running Weight : 255
```

Forwarder 01

```
State : Active
Virtual MAC : 000f-e2ff-0011 (Take Over)
Owner ID : 0000-5e01-1101
Priority : 85
Active : local
Redirect Time : 93 secs
Time-out Time : 1293 secs
```

Forwarder 02

```
State : Listening
Virtual MAC : 000f-e2ff-0012 (Learnt)
Owner ID : 0000-5e01-1103
Priority : 85
Active : 10.1.1.3
```

Forwarder 03

```
State : Active
Virtual MAC : 000f-e2ff-0013 (Owner)
Owner ID : 0000-5e01-1105
Priority : 255
Active : local
```

Forwarder Weight Track Information:

```

Track Object : 1 State : Positive Weight Reduced : 250
Forwarder Switchover Track Information:
Track Object : 2 State : Positive
 Member IP : 10.1.1.2
Track Object : 3 State : Positive
 Member IP : 10.1.1.3

```

The output shows that when VLAN interface 3 on Switch A fails, the weight of the AVF on Switch A decreases to 5, which is lower than the lower limit of failure. All VFs on Switch A turn to initialized state and cannot be used for packet forwarding. The VF corresponding to MAC address 000f-e2ff-0011 on Switch C becomes the AVF, and Switch C takes over Switch A for packet forwarding.

# When the timeout timer (about 1800 seconds) expires, display the detailed information about VRRP group 1 on Switch C.

```

[SwitchC-Vlan-interface2] display vrrp verbose
IPv4 Standby Information:
 Run Mode : Load Balance
 Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 1
 Admin Status : Up State : Backup
 Config Pri : 100 Running Pri : 100
 Preempt Mode : Yes Delay Time : 5
 Become Master : 4200ms left
 Auth Type : None
 Virtual IP : 10.1.1.1
 Member IP List : 10.1.1.4 (Local, Backup)
 10.1.1.2 (Master)
 10.1.1.3 (Backup)
Forwarder Information: 2 Forwarders 1 Active
 Config Weight : 255
 Running Weight : 255
Forwarder 02
 State : Listening
 Virtual MAC : 000f-e2ff-0012 (Learnt)
 Owner ID : 0000-5e01-1103
 Priority : 127
 Active : 10.1.1.3
Forwarder 03
 State : Active
 Virtual MAC : 000f-e2ff-0013 (Owner)
 Owner ID : 0000-5e01-1105
 Priority : 255
 Active : local
Forwarder Weight Track Information:
 Track Object : 1 State : Positive Weight Reduced : 250
Forwarder Switchover Track Information:
 Track Object : 2 State : Positive

```

```
Member IP : 10.1.1.2
Track Object : 3 State : Positive
Member IP : 10.1.1.3
```

The output shows that when the timeout timer expires, the VF corresponding to virtual MAC address 000f-e2ff-0011 is removed, and does not forward the packets destined for the MAC address any more.

# When Switch A fails, use the **display vrrp verbose** command to display the detailed information about VRRP group 1 on Switch B.

```
[SwitchB-Vlan-interface2] display vrrp verbose
IPv4 Standby Information:
 Run Mode : Load Balance
 Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 1
 Admin Status : Up State : Master
 Config Pri : 110 Running Pri : 110
 Preempt Mode : Yes Delay Time : 5
 Auth Type : None
 Virtual IP : 10.1.1.1
 Member IP List : 10.1.1.3 (Local, Master)
 10.1.1.4 (Backup)
Forwarder Information: 2 Forwarders 1 Active
 Config Weight : 255
 Running Weight : 255
Forwarder 02
 State : Active
 Virtual MAC : 000f-e2ff-0012 (Owner)
 Owner ID : 0000-5e01-1103
 Priority : 255
 Active : local
Forwarder 03
 State : Listening
 Virtual MAC : 000f-e2ff-0013 (Learnt)
 Owner ID : 0000-5e01-1105
 Priority : 127
 Active : 10.1.1.4
Forwarder Weight Track Information:
 Track Object : 1 State : Positive Weight Reduced : 250
```

The output shows that when Switch A fails, Switch B becomes the master because its priority is higher than that of Switch C.

# When Switch B fails, use the **display vrrp verbose** command to display the detailed information about VRRP group 1 on Switch C.

```
[SwitchC-Vlan-interface2] display vrrp verbose
IPv4 Standby Information:
 Run Mode : Load Balance
 Run Method : Virtual MAC
Total number of virtual routers : 1
```

```

Interface GigabitEthernet1/0/1
 VRID : 1 Adver Timer : 1
 Admin Status : Up State : Master
 Config Pri : 100 Running Pri : 100
 Preempt Mode : Yes Delay Time : 5
 Auth Type : None
 Virtual IP : 10.1.1.1
 Member IP List : 10.1.1.4 (Local, Master)
Forwarder Information: 2 Forwarders 2 Active
 Config Weight : 255
 Running Weight : 255
Forwarder 02
 State : Active
 Virtual MAC : 000f-e2ff-0012 (Take Over)
 Owner ID : 0000-5e01-1103
 Priority : 85
 Active : local
 Redirect Time : 93 secs
 Time-out Time : 1293 secs
Forwarder 03
 State : Active
 Virtual MAC : 000f-e2ff-0013 (Owner)
 Owner ID : 0000-5e01-1105
 Priority : 255
 Active : local
Forwarder Weight Track Information:
 Track Object : 1 State : Positive Weight Reduced : 250
Forwarder Switchover Track Information:
 Track Object : 2 State : Negative
 Member IP : 10.1.1.2
 Track Object : 3 State : Negative
 Member IP : 10.1.1.3

```

The output shows that when Switch B fails, Switch C becomes the master, and Forwarder 02 on Switch C immediately becomes active. Switch C takes over the AVF on Switch B.

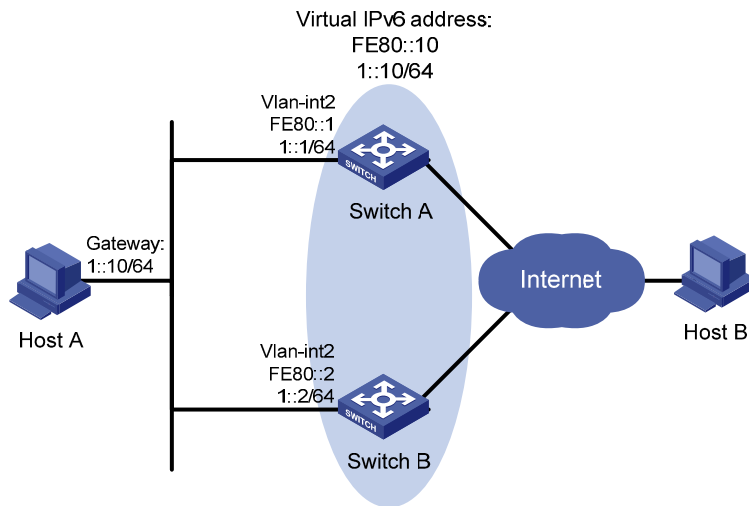
## IPv6-based VRRP configuration examples

### Single VRRP group configuration example

#### Network requirements

- Switch A and Switch B belong to VRRP group 1 with the virtual IP addresses of 1::10/64 and FE80::10.
- Host A wants to access Host B on the Internet, and learns 1::10/64 as its default gateway through RA messages sent by the switches.
- When Switch A operates normally, packets sent from Host A to Host B are forwarded by Switch A; when Switch A fails, packets sent from Host A to Host B are forwarded by Switch B.

Figure 44 Network diagram



## Configuration procedure

### 1. Configure Switch A:

# Configure VLAN 2.

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] vlan 2
[SwitchA-vlan2] port gigabitethernet 1/0/5
[SwitchA-vlan2] quit
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ipv6 address fe80::1 link-local
[SwitchA-Vlan-interface2] ipv6 address 1::1 64
```

# Create a VRRP group 1 and set its virtual IPv6 addresses to FE80::10 and 1::10.

```
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 virtual-ip fe80::10 link-local
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 virtual-ip 1::10
```

# Set the priority of Switch A in VRRP group 1 to 110, which is higher than that of Switch B (100), so that Switch A can become the master.

```
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 priority 110
```

# Configure Switch A to operate in preemptive mode so that it can become the master whenever it works normally, and configure the preemption delay as five seconds to avoid frequent status switchover.

```
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 preempt-mode timer delay 5
```

# Enable Switch A to send RA messages, so that Host A can learn the default gateway address.

```
[SwitchA-Vlan-interface2] undo ipv6 nd ra halt
```

### 2. Configure Switch B:

# Configure VLAN 2.

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] vlan 2
[SwitchB-vlan2] port gigabitethernet 1/0/5
[SwitchB-vlan2] quit
```

```
[SwitchB] interface vlan-interface 2
[SwitchB-Vlan-interface2] ipv6 address fe80::2 link-local
[SwitchB-Vlan-interface2] ipv6 address 1::2 64
Create a VRRP group 1 and set its virtual IPv6 addresses to FE80::10 and 1::10.
[SwitchB-Vlan-interface2] vrrp ipv6 vrid 1 virtual-ip fe80::10 link-local
[SwitchB-Vlan-interface2] vrrp ipv6 vrid 1 virtual-ip 1::10
Configure Switch B to operate in preemptive mode, with the preemption delay set to 5 seconds.
[SwitchB-Vlan-interface2] vrrp ipv6 vrid 1 preempt-mode timer delay 5
Enable Switch B to send RA messages, so that Host A can learn the default gateway address.
[SwitchB-Vlan-interface2] vrrp ipv6 vrid 1 preempt-mode timer delay 5
```

### 3. Verify the configuration:

After the configuration, Host B can be pinged successfully on Host A. To verify your configuration, use the **display vrrp ipv6 verbose** command.

# Display the detailed information about VRRP group 1 on Switch A.

```
[SwitchA-Vlan-interface2] display vrrp ipv6 verbose
IPv6 Standby Information:
 Run Mode : Standard
 Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 100
 Admin Status : Up State : Master
 Config Pri : 110 Running Pri : 110
 Preempt Mode : Yes Delay Time : 5
 Auth Type : None
 Virtual IP : FE80::10
 1::10
 Virtual MAC : 0000-5e00-0201
 Master IP : FE80::1
```

# Display the detailed information about VRRP group 1 on Switch B.

```
[SwitchB-Vlan-interface2] display vrrp ipv6 verbose
IPv6 Standby Information:
 Run Mode : Standard
 Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 100
 Admin Status : Up State : Backup
 Config Pri : 100 Running Pri : 100
 Preempt Mode : Yes Delay Time : 5
 Become Master : 4200ms left
 Auth Type : None
 Virtual IP : FE80::10
 1::10
 Master IP : FE80::1
```

The output shows that in VRRP group 1 Switch A is the master, Switch B is the backup and packets sent from Host A to Host B are forwarded by Switch A.



When Switch A fails, you can still successfully ping Host B on Host A. To view the detailed information about the VRRP group on Switch B, use the **display vrrp ipv6 verbose** command.

# When Switch A fails, the detailed information about VRRP group 1 on Switch B is displayed.

```
[SwitchB-Vlan-interface2] display vrrp ipv6 verbose
IPv6 Standby Information:
 Run Mode : Standard
 Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 100
 Admin Status : Up State : Master
 Config Pri : 100 Running Pri : 100
 Preempt Mode : Yes Delay Time : 5
 Auth Type : None
 Virtual IP : FE80::10
 1::10
 Virtual MAC : 0000-5e00-0201
 Master IP : FE80::2
```

The output shows that when Switch A fails, Switch B becomes the master, and packets sent from Host A to Host B are forwarded by Switch B.

# After Switch A resumes normal operation, use the **display vrrp ipv6 verbose** command to display the detailed information about VRRP group 1 on Switch A.

```
[SwitchA-Vlan-interface2] display vrrp ipv6 verbose
IPv6 Standby Information:
 Run Mode : Standard
 Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 100
 Admin Status : Up State : Master
 Config Pri : 110 Running Pri : 110
 Preempt Mode : Yes Delay Time : 5
 Auth Type : None
 Virtual IP : FE80::10
 1::10
 Virtual MAC : 0000-5e00-0201
 Master IP : FE80::1
```

The output shows that after Switch A resumes normal operation, it becomes the master, and packets sent from host A to host B are forwarded by Switch A.

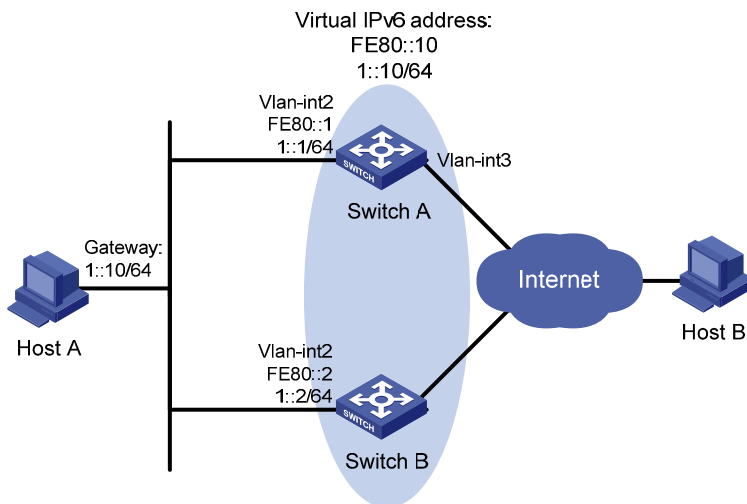
## VRRP interface tracking configuration example

### Network requirements

- Switch A and Switch B belong to VRRP group 1 with the virtual IP addresses of 1::10/64 and FE80::10.
- Host A wants to access Host B on the Internet, and learns 1::10/64 as its default gateway through RA messages sent by the switches.

- When Switch A operates normally, packets sent from Host A to Host B are forwarded by Switch A. If VLAN-interface 3 through which Switch A connects to the Internet is not available, packets sent from Host A to Host B are forwarded by Switch B.
- To prevent attacks to the VRRP group from illegal users by using spoofed packets, configure the authentication mode as plain text to authenticate the VRRP packets in VRRP group 1, and specify the authentication key as **hello**.

**Figure 45 Network diagram**



## Configuration procedure

### 1. Configure Switch A:

# Configure VLAN 2.

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] vlan 2
[SwitchA-vlan2] port gigabitethernet 1/0/5
[SwitchA-vlan2] quit
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ipv6 address fe80::1 link-local
[SwitchA-Vlan-interface2] ipv6 address 1::1 64
```

# Create a VRRP group 1 and set its virtual IPv6 addresses to FE80::10 and 1::10.

```
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 virtual-ip fe80::10 link-local
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 virtual-ip 1::10
```

# Set the priority of Switch A in VRRP group 1 to 110, which is higher than that of Switch B (100), so that Switch A can become the master.

```
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 priority 110
```

# Set the authentication mode for VRRP group 1 to **simple** and authentication key to **hello**.

```
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 authentication-mode simple hello
```

# Set the VRRP advertisement interval to 400 centiseconds.

```
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 timer advertise 400
```

# Configure Switch A to operate in preemptive mode, so that it can become the master whenever it works normally; configure the preemption delay as five seconds to avoid frequent status switchover.

```
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 preempt-mode timer delay 5
Set VLAN-interface 3 on Switch A to be tracked, and configure the amount by which the priority
value decreases to be more than 10 (30 in this example), so that when VLAN interface 3 fails, the
priority of Switch A in VRRP group 1 decreases to a value lower than 100 and thus Switch B can
become the master.
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 track interface vlan-interface 3 reduced
30
Enable Switch A to send RA messages, so that Host A can learn the default gateway address.
[SwitchA-Vlan-interface2] undo ipv6 nd ra halt
```

## 2. Configure Switch B:

# Configure VLAN 2.

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB] vlan 2
[SwitchB-vlan2] port gigabitethernet 1/0/5
[SwitchB-vlan2] quit
[SwitchB] interface vlan-interface 2
[SwitchB-Vlan-interface2] ipv6 address fe80::2 link-local
[SwitchB-Vlan-interface2] ipv6 address 1::2 64
```

# Create a VRRP group 1 and set its virtual IPv6 addresses to FE80::10 and 1::10.

```
[SwitchB-Vlan-interface2] vrrp ipv6 vrid 1 virtual-ip fe80::10 link-local
[SwitchB-Vlan-interface2] vrrp ipv6 vrid 1 virtual-ip 1::10
```

# Set the authentication mode for VRRP group 1 to **simple** and authentication key to **hello**.

```
[SwitchB-Vlan-interface2] vrrp ipv6 vrid 1 authentication-mode simple hello
```

# Set the VRRP advertisement interval to 400 centiseconds.

```
[SwitchB-Vlan-interface2] vrrp ipv6 vrid 1 timer advertise 400
```

# Configure Switch B to operate in preemptive mode, so that Switch B can become the master after the priority of Switch A decreases to a value lower than 100. Configure the preemption delay as five seconds to avoid frequent status switchover.

```
[SwitchB-Vlan-interface2] vrrp ipv6 vrid 1 preempt-mode timer delay 5
```

# Enable Switch B to send RA messages, so that Host A can learn the default gateway address.

```
[SwitchB-Vlan-interface2] undo ipv6 nd ra halt
```

## 3. Verify the configuration:

After the configuration, Host B can be pinged successfully on Host A. To verify the configuration, use the **display vrrp ipv6 verbose** command.

# Display the detailed information about VRRP group 1 on Switch A.

```
[SwitchA-Vlan-interface2] display vrrp ipv6 verbose
IPv6 Standby Information:
 Run Mode : Standard
 Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 400
 Admin Status : Up State : Master
 Config Pri : 110 Running Pri : 110
 Preempt Mode : Yes Delay Time : 5
```

```

Auth Type : Simple Key : *****
Virtual IP : FE80::10
 1::10
Virtual MAC : 0000-5e00-0201
Master IP : FE80::1

```

VRRP Track Information:

```

Track Interface: Vlan3 State : Up Pri Reduced : 30

```

# Display the detailed information about VRRP group 1 on Switch B.

```
[SwitchB-Vlan-interface2] display vrrp ipv6 verbose
```

IPv6 Standby Information:

```

Run Mode : Standard
Run Method : Virtual MAC

```

Total number of virtual routers : 1

Interface Vlan-interface2

```

VRID : 1 Adver Timer : 400
Admin Status : Up State : Backup
Config Pri : 100 Running Pri : 100
Preempt Mode : Yes Delay Time : 5
Become Master : 4200ms left
Auth Type : Simple Key : *****
Virtual IP : FE80::10
 1::10
Master IP : FE80::1

```

The output shows that in VRRP group 1 Switch A is the master, Switch B is the backup and packets sent from Host A to Host B are forwarded by Switch A.

When interface VLAN-interface 3 on Switch A is not available, you can still ping Host B successfully on Host A. To view the detailed information about the VRRP group, use the **display vrrp ipv6 verbose** command.

# When interface VLAN-interface 3 on Switch A is not available, the detailed information about VRRP group 1 on Switch A is displayed.

```
[SwitchA-Vlan-interface2] display vrrp ipv6 verbose
```

IPv6 Standby Information:

```

Run Mode : Standard
Run Method : Virtual MAC

```

Total number of virtual routers : 1

Interface Vlan-interface2

```

VRID : 1 Adver Timer : 400
Admin Status : Up State : Backup
Config Pri : 110 Running Pri : 80
Preempt Mode : Yes Delay Time : 5
Become Master : 4200ms left
Auth Type : Simple Key : *****
Virtual IP : FE80::10
 1::10
Master IP : FE80::2

```

VRRP Track Information:

```

Track Interface: Vlan3 State : Down Pri Reduced : 30

```

# When interface VLAN-interface 3 on Switch A is not available, the detailed information about VRRP group 1 on Switch B is displayed.

```
[SwitchB-Vlan-interface2] display vrrp ipv6 verbose
IPv6 Standby Information:
 Run Mode : Standard
 Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 400
 Admin Status : Up State : Master
 Config Pri : 100 Running Pri : 100
 Preempt Mode : Yes Delay Time : 5
 Auth Type : Simple Key : *****
 Virtual IP : FE80::10
 1::10
 Virtual MAC : 0000-5e00-0201
 Master IP : FE80::2
```

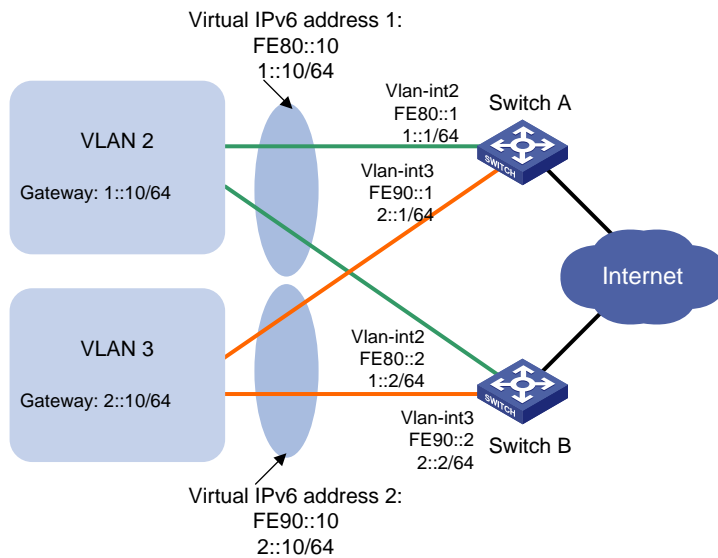
The output shows that when VLAN-interface 3 on Switch A is not available, the priority of Switch A is reduced to 80 and Switch A becomes the backup. Switch B becomes the master and packets sent from Host A to Host B are forwarded by Switch B.

## VRRP with multiple VLANs configuration example

### Network requirements

- Switch A and Switch B belong to both VRRP group 1 and VRRP group 2. The virtual IPv6 addresses of VRRP group 1 are 1::10/64 and FE80::10, and those of VRRP group 2 are 2::10/64 and FE90::10.
- Hosts in VLAN 2 learn 1::10/64 as their default gateway and hosts in VLAN 3 learn 2::10/64 as their default gateway through RA messages sent by the switches.
- In VRRP group 1, Switch A has a higher priority than Switch B. In VRRP group 2, Switch B has a higher priority than Switch A. In this case, hosts in VLAN 2 and VLAN 3 can communicate with external networks through Switch A and Switch B, respectively, and when Switch A or Switch B fails, the hosts can use the other switch to communicate with external networks to avoid communication interruption.

**Figure 46 Network diagram**



## Configuration procedure

### 1. Configure Switch A:

# Configure VLAN 2.

```
<SwitchA> system-view
[SwitchA] ipv6
[SwitchA] vlan 2
[SwitchA-vlan2] port gigabitethernet 1/0/5
[SwitchA-vlan2] quit
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ipv6 address fe80::1 link-local
[SwitchA-Vlan-interface2] ipv6 address 1::1 64
```

# Create VRRP group 1 and set its virtual IPv6 addresses to FE80::10 to 1::10.

```
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 virtual-ip fe80::10 link-local
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 virtual-ip 1::10
```

# Set the priority of Switch A in VRRP group 1 to 110, which is higher than that of Switch B (100), so that Switch A can become the master in VRRP group 1.

```
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 priority 110
[SwitchA-Vlan-interface2] quit
```

# Enable Switch A to send RA messages, so that hosts in VLAN 2 can learn the default gateway address.

```
[SwitchA-Vlan-interface2] undo ipv6 nd ra halt
[SwitchA-Vlan-interface2] quit
```

# Configure VLAN 3.

```
[SwitchA] vlan 3
[SwitchA-vlan3] port gigabitethernet 1/0/6
[SwitchA-vlan3] quit
[SwitchA] interface vlan-interface 3
[SwitchA-Vlan-interface3] ipv6 address fe90::1 link-local
[SwitchA-Vlan-interface3] ipv6 address 2::1 64
```

```
Create VRRP group 2 and set its virtual IPv6 addresses to FE90::10 and 2::10.
[SwitchA-Vlan-interface3] vrrp ipv6 vrid 2 virtual-ip fe90::10 link-local
[SwitchA-Vlan-interface3] vrrp ipv6 vrid 2 virtual-ip 2::10
Enable Switch A to send RA messages, so that hosts in VLAN 3 can learn the default gateway
address.
[SwitchA-Vlan-interface2] undo ipv6 nd ra halt
```

## 2. Configure Switch B:

```
Configure VLAN 2.
```

```
<SwitchB> system-view
[SwitchB] ipv6
[SwitchB-vlan2] port gigabitethernet 1/0/5
[SwitchB-vlan2] quit
[SwitchB] interface vlan-interface 2
[SwitchB-Vlan-interface2] ipv6 address fe80::2 link-local
[SwitchB-Vlan-interface2] ipv6 address 1::2 64
```

```
Create VRRP group 1 and set its virtual IPv6 addresses to FE80::10 and 1::10.
```

```
[SwitchB-Vlan-interface2] vrrp ipv6 vrid 1 virtual-ip fe80::10 link-local
[SwitchB-Vlan-interface2] vrrp ipv6 vrid 1 virtual-ip 1::10
[SwitchB-Vlan-interface2] quit
```

```
Enable Switch B to send RA messages, so that hosts in VLAN 2 can learn the default gateway
address.
```

```
[SwitchB-Vlan-interface2] undo ipv6 nd ra halt
[SwitchB-Vlan-interface2] quit
```

```
Configure VLAN 3.
```

```
[SwitchB] vlan 3
[SwitchB-vlan3] port gigabitethernet 1/0/6
[SwitchB-vlan3] quit
[SwitchB] interface vlan-interface 3
[SwitchB-Vlan-interface3] ipv6 address fe90::2 link-local
[SwitchB-Vlan-interface3] ipv6 address 2::2 64
```

```
Create VRRP group 2 and set its virtual IPv6 addresses to FE90::10 and 2::10.
```

```
[SwitchB-Vlan-interface3] vrrp ipv6 vrid 2 virtual-ip fe90::10 link-local
[SwitchB-Vlan-interface3] vrrp ipv6 vrid 2 virtual-ip 2::10
```

```
Set the priority of Switch B in VRRP group 2 to 110, which is higher than that of Switch A (100),
so that Switch B can become the master in VRRP group 2.
```

```
[SwitchB-Vlan-interface3] vrrp ipv6 vrid 2 priority 110
```

```
Enable Switch B to send RA messages, so that hosts in VLAN 3 can learn the default gateway
address.
```

```
[SwitchB-Vlan-interface3] undo ipv6 nd ra halt
```

## 3. Verify the configuration:

To verify the configuration, use the **display vrrp ipv6 verbose** command.

```
Display the detailed information about the VRRP group on Switch A.
```

```
[SwitchA-Vlan-interface3] display vrrp ipv6 verbose
IPv6 Standby Information:
 Run Mode : Standard
 Run Method : Virtual MAC
```

Total number of virtual routers : 2

Interface Vlan-interface2

VRID	: 1	Adver Timer	: 100
Admin Status	: Up	State	: Master
Config Pri	: 110	Running Pri	: 110
Preempt Mode	: Yes	Delay Time	: 0
Auth Type	: None		
Virtual IP	: FE80::10		
	1::10		
Virtual MAC	: 0000-5e00-0201		
Master IP	: FE80::1		

Interface Vlan-interface3

VRID	: 2	Adver Timer	: 100
Admin Status	: Up	State	: Backup
Config Pri	: 100	Running Pri	: 100
Preempt Mode	: Yes	Delay Time	: 0
Become Master	: 2200ms left		
Auth Type	: None		
Virtual IP	: FE90::10		
	2::10		
Master IP	: FE90::2		

# Display the detailed information about the VRRP group on Switch B.

[SwitchB-Vlan-interface3] display vrrp ipv6 verbose

IPv6 Standby Information:

Run Mode	: Standard
Run Method	: Virtual MAC

Total number of virtual routers : 2

Interface Vlan-interface2

VRID	: 1	Adver Timer	: 100
Admin Status	: Up	State	: Backup
Config Pri	: 100	Running Pri	: 100
Preempt Mode	: Yes	Delay Time	: 0
Become Master	: 2200ms left		
Auth Type	: None		
Virtual IP	: FE80::10		
	1::10		
Master IP	: FE80::1		

Interface Vlan-interface3

VRID	: 2	Adver Timer	: 100
Admin Status	: Up	State	: Master
Config Pri	: 110	Running Pri	: 110
Preempt Mode	: Yes	Delay Time	: 0
Auth Type	: None		
Virtual IP	: FE90::10		
	2::10		
Virtual MAC	: 0000-5e00-0202		
Master IP	: FE90::2		



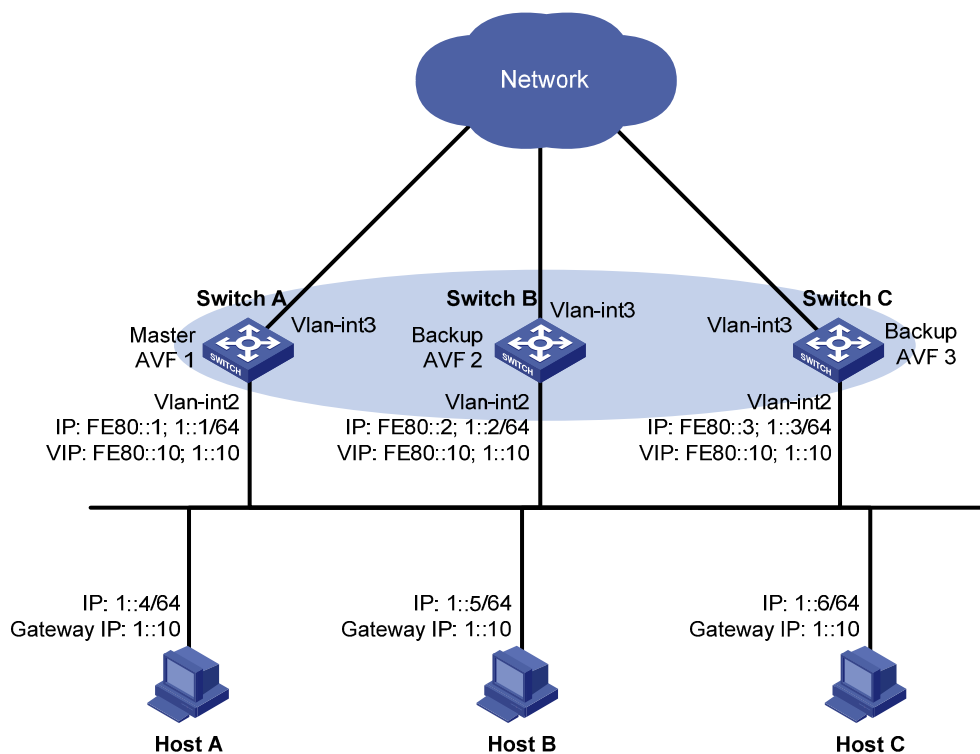
The output shows that in VRRP group 1 Switch A is the master, Switch B is the backup and hosts with the default gateway of 1::10/64 accesses the Internet through Switch A; in VRRP group 2 Switch A is the backup, Switch B is the master and hosts with the default gateway of 2::10/64 accesses the Internet through Switch B.

## VRRP load balancing mode configuration example

### Network requirements

- Switch A, Switch B, and Switch C belong to VRRP group 1 with the virtual IPv6 addresses of FE80::10 and 1::10.
- Hosts on network segment 1::/64 learn 1::10 as their default gateway through RA messages sent by the switches. Use the VRRP group to make sure that when a gateway (Switch A, Switch B, or Switch C) fails, the hosts on the LAN can access the external network through another gateway.
- VRRP group 1 is operating in load balancing mode to make good use of network resources.
- Configure a track entry on Switch A, Switch B, and Switch C to monitor their own VLAN-interface 3. When the interface on Switch A, Switch B, or Switch C fails, the weight of the corresponding switch decreases so that another switch with a higher weight can take over.

Figure 47 Network diagram



### Configuration procedure

1. Configure Switch A:  

```
Configure VLAN 2.
<SwitchA> system-view
[SwitchA] vlan 2
[SwitchA-vlan2] port gigabitethernet 1/0/5
[SwitchA-vlan2] quit
```

```

Configure VRRP to operate in load balancing mode.
[SwitchA] vrrp mode load-balance
Create VRRP group 1 and configure its virtual IPv6 addresses as FE80::10 and 1::10.
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ipv6 address fe80::1 link-local
[SwitchA-Vlan-interface2] ipv6 address 1::1 64
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 virtual-ip fe80::10 link-local
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 virtual-ip 1::10
Set the priority of Switch A in VRRP group 1 to 120, which is higher than that of Switch B (110)
and that of Switch C (100), so that Switch A can become the master.
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 priority 120
Configure Switch A to operate in preemptive mode so that it can become the master whenever
it works normally; configure the preemption delay as five seconds to avoid frequent status
switchover.
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 preempt-mode timer delay 5
Enable Switch A to send RA messages so that hosts on network segment 1::/64 can learn the
default gateway address.
[SwitchA-Vlan-interface2] undo ipv6 nd ra halt
[SwitchA-Vlan-interface2] quit
Create track entry 1 to associate with the physical status of VLAN-interface 3 on Switch A. When
the track entry becomes negative, it means that the interface fails.
[SwitchA] track 1 interface vlan-interface 3
Configure VF tracking to monitor track entry 1 and specify the value by which the weight
decreases, making the weight of Switch A decrease by more than 245 (250 in this example) when
track entry 1 turns to negative. In such a case, another router with a higher weight can take over.
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] vrrp ipv6 vrid 1 weight track 1 reduced 250

```

## 2. Configure Switch B:

```

Configure VLAN 2.
<SwitchB> system-view
[SwitchB] vlan 2
[SwitchB-vlan2] port gigabitethernet 1/0/5
[SwitchB-vlan2] quit
Configure VRRP to operate in load balancing mode.
[SwitchB] vrrp mode load-balance
Create VRRP group 1 and configure its virtual IPv6 addresses as FE80::10 and 1::10.
[SwitchB] interface vlan-interface 2
[SwitchB-Vlan-interface2] ipv6 address fe80::2 link-local
[SwitchB-Vlan-interface2] ipv6 address 1::2 64
[SwitchB-Vlan-interface2] vrrp ipv6 vrid 1 virtual-ip fe80::10 link-local
[SwitchB-Vlan-interface2] vrrp ipv6 vrid 1 virtual-ip 1::10
Set the priority of Switch B in VRRP group 1 to 110, which is higher than that of Switch C (100)
so that Switch B can become the master when Switch A fails.
[SwitchB-Vlan-interface2] vrrp ipv6 vrid 1 priority 110
Set Switch B to operate in preemptive mode and set the preemption delay to five seconds.
[SwitchB-Vlan-interface2] vrrp ipv6 vrid 1 preempt-mode timer delay 5

```

# Enable Switch B to send RA messages so that hosts on network segment 1::/64 can learn the default gateway address.

```
[SwitchB-Vlan-interface2] undo ipv6 nd ra halt
[SwitchB-Vlan-interface2] quit
```

# Create track entry 1 to associate with the physical status of VLAN-interface 3 on Switch B. When the track entry becomes negative, it means that the interface fails.

```
[SwitchB] track 1 interface vlan-interface 3
```

# Configure VF tracking to monitor track entry 1 and specify the value by which the weight decreases, making the weight of Switch B decrease by more than 245 (250 in this example) when track entry 1 turns to negative. In such a case, another router with a higher weight can take over.

```
[SwitchB] interface vlan-interface 2
```

```
[SwitchB-Vlan-interface2] vrrp ipv6 vrid 1 weight track 1 reduced 250
```

### 3. Configure Switch C:

# Configure VLAN 2.

```
<SwitchC> system-view
```

```
[SwitchC] vlan 2
```

```
[SwitchC-vlan2] port gigabitethernet 1/0/5
```

```
[SwitchC-vlan2] quit
```

# Configure VRRP to operate in load balancing mode.

```
[SwitchC] vrrp mode load-balance
```

# Create VRRP group 1 and configure its virtual IPv6 addresses as FE80::10 and 1::10.

```
[SwitchC] interface vlan-interface 2
```

```
[SwitchC-Vlan-interface2] ipv6 address fe80::3 link-local
```

```
[SwitchC-Vlan-interface2] ipv6 address 1::3 64
```

```
[SwitchC-Vlan-interface2] vrrp ipv6 vrid 1 virtual-ip fe80::10 link-local
```

```
[SwitchC-Vlan-interface2] vrrp ipv6 vrid 1 virtual-ip 1::10
```

# Set Switch C to operate in preemptive mode and set the preemption delay to five seconds.

```
[SwitchC-Vlan-interface2] vrrp ipv6 vrid 1 preempt-mode timer delay 5
```

# Enable Switch C to send RA messages, so that hosts on network segment 1::/64 can learn the default gateway address.

```
[SwitchC-Vlan-interface2] undo ipv6 nd ra halt
```

```
[SwitchC-Vlan-interface2] quit
```

# Create track entry 1 to associate with the physical status of VLAN-interface 3 on Switch C. When the track entry becomes negative, it means that the interface fails.

```
[SwitchC] track 1 interface vlan-interface 3
```

# Configure VF tracking to monitor track entry 1 and specify the value by which the weight decreases, making the weight of Switch C decrease by more than 245 (250 in this example) when track entry 1 turns to negative. In such a case, another router with a higher weight can take over.

```
[SwitchC] interface vlan-interface 2
```

```
[SwitchC-Vlan-interface2] vrrp ipv6 vrid 1 weight track 1 reduced 250
```

### 4. Verify the configuration:

After the configuration, Host A can ping the external network. To verify the configuration, use the **display ipv6 vrrp verbose** command.

# Display the detailed information about VRRP group 1 on Switch A.

```
[SwitchA-Vlan-interface2] display vrrp ipv6 verbose
```

```
IPv6 Standby Information:
```

```

Run Mode : Load Balance
Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 100
 Admin Status : Up State : Master
 Config Pri : 120 Running Pri : 120
 Preempt Mode : Yes Delay Time : 5
 Auth Type : None
 Virtual IP : FE80::10
 1::10
 Member IP List : FE80::1 (Local, Master)
 FE80::2 (Backup)
 FE80::3 (Backup)
Forwarder Information: 3 Forwarders 1 Active
 Config Weight : 255
 Running Weight : 255
Forwarder 01
 State : Active
 Virtual MAC : 000f-e2ff-4011 (Owner)
 Owner ID : 0000-5e01-1101
 Priority : 255
 Active : local
Forwarder 02
 State : Listening
 Virtual MAC : 000f-e2ff-4012 (Learnt)
 Owner ID : 0000-5e01-1103
 Priority : 127
 Active : FE80::2
Forwarder 03
 State : Listening
 Virtual MAC : 000f-e2ff-4013 (Learnt)
 Owner ID : 0000-5e01-1105
 Priority : 127
 Active : FE80::3
Forwarder Weight Track Information:
 Track Object : 1 State : Positive Weight Reduced : 250

```

# Display the detailed information about VRRP group 1 on Switch B.

```
[SwitchB-Vlan-interface2] display vrrp ipv6 verbose
```

```

IPv6 Standby Information:
 Run Mode : Load Balance
 Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
 VRID : 1 Adver Timer : 100
 Admin Status : Up State : Backup
 Config Pri : 110 Running Pri : 110
 Preempt Mode : Yes Delay Time : 5

```

```

Become Master : 2200ms left
Auth Type : None
Virtual IP : FE80::10
 1::10
Member IP List : FE80::2 (Local, Backup)
 FE80::1 (Master)
 FE80::3 (Backup)
Forwarder Information: 3 Forwarders 1 Active
 Config Weight : 255
 Running Weight : 255
Forwarder 01
 State : Listening
 Virtual MAC : 000f-e2ff-4011 (Learnt)
 Owner ID : 0000-5e01-1101
 Priority : 127
 Active : FE80::1
Forwarder 02
 State : Active
 Virtual MAC : 000f-e2ff-4012 (Owner)
 Owner ID : 0000-5e01-1103
 Priority : 255
 Active : local
Forwarder 03
 State : Listening
 Virtual MAC : 000f-e2ff-4013 (Learnt)
 Owner ID : 0000-5e01-1105
 Priority : 127
 Active : FE80::3
Forwarder Weight Track Information:
 Track Object : 1 State : Positive Weight Reduced : 250

```

# Display the detailed information about VRRP group 1 on Switch C.

```
[SwitchC-Vlan-interface2] display vrrp ipv6 verbose
```

```
IPv6 Standby Information:
```

```
Run Mode : Load Balance
Run Method : Virtual MAC
```

```
Total number of virtual routers : 1
```

```
Interface Vlan-interface2
```

```

VRID : 1 Adver Timer : 100
Admin Status : Up State : Backup
Config Pri : 100 Running Pri : 100
Preempt Mode : Yes Delay Time : 5
Become Master : 4200ms left
Auth Type : None
Virtual IP : FE80::10
 1::10
Member IP List : FE80::3 (Local, Backup)
 FE80::1 (Master)
 FE80::2 (Backup)

```

Forwarder Information: 3 Forwarders 1 Active

Config Weight : 255  
Running Weight : 255

Forwarder 01

State : Listening  
Virtual MAC : 000f-e2ff-4011 (Learnt)  
Owner ID : 0000-5e01-1101  
Priority : 127  
Active : FE80::1

Forwarder 02

State : Listening  
Virtual MAC : 000f-e2ff-4012 (Learnt)  
Owner ID : 0000-5e01-1103  
Priority : 127  
Active : FE80::2

Forwarder 03

State : Active  
Virtual MAC : 000f-e2ff-4013 (Owner)  
Owner ID : 0000-5e01-1105  
Priority : 255  
Active : local

Forwarder Weight Track Information:

Track Object : 1 State : Positive Weight Reduced : 250

The output shows that in VRRP group 1, Switch A is the master and Switch B and Switch C are the backups. Each switch has one AVF and two LVFs acting as the backups.

# When VLAN interface 3 on Switch A fails, use the **display vrrp ipv6 verbose** command to display the detailed information about VRRP group 1 on Switch A.

```
[SwitchA-Vlan-interface2] display vrrp ipv6 verbose
```

IPv6 Standby Information:

Run Mode : Load Balance  
Run Method : Virtual MAC

Total number of virtual routers : 1

Interface Vlan-interface2

VRID : 1 Adver Timer : 100  
Admin Status : Up State : Master  
Config Pri : 120 Running Pri : 120  
Preempt Mode : Yes Delay Time : 5  
Auth Type : None  
Virtual IP : FE80::10  
1::10  
Member IP List : FE80::1 (Local, Master)  
FE80::2 (Backup)  
FE80::3 (Backup)

Forwarder Information: 3 Forwarders 0 Active

Config Weight : 255  
Running Weight : 5

Forwarder 01

State : Initialize

```
Virtual MAC : 000f-e2ff-4011 (Owner)
Owner ID : 0000-5e01-1101
Priority : 0
Active : FE80::3
```

Forwarder 02

```
State : Initialize
Virtual MAC : 000f-e2ff-4012 (Learnt)
Owner ID : 0000-5e01-1103
Priority : 0
Active : FE80::2
```

Forwarder 03

```
State : Initialize
Virtual MAC : 000f-e2ff-4013 (Learnt)
Owner ID : 0000-5e01-1105
Priority : 0
Active : FE80::3
```

Forwarder Weight Track Information:

```
Track Object : 1 State : Negative Weight Reduced : 250
```

# Use the **display vrrp ipv6 verbose** command to display the detailed information about VRRP group 1 on Switch C.

```
[SwitchC-Vlan-interface2] display vrrp ipv6 verbose
```

IPv6 Standby Information:

```
Run Mode : Load Balance
Run Method : Virtual MAC
```

Total number of virtual routers : 1

Interface Vlan-interface2

```
VRID : 1 Adver Timer : 100
Admin Status : Up State : Backup
Config Pri : 100 Running Pri : 100
Preempt Mode : Yes Delay Time : 5
Become Master : 4200ms left
Auth Type : None
Virtual IP : FE80::10
 1::10
Member IP List : FE80::3 (Local, Backup)
 FE80::1 (Master)
 FE80::2 (Backup)
```

Forwarder Information: 3 Forwarders 2 Active

```
Config Weight : 255
Running Weight : 255
```

Forwarder 01

```
State : Active
Virtual MAC : 000f-e2ff-4011 (Take Over)
Owner ID : 0000-5e01-1101
Priority : 85
Active : local
Redirect Time : 93 secs
Time-out Time : 1293 secs
```

Forwarder 02

State : Listening  
Virtual MAC : 000f-e2ff-4012 (Learnt)  
Owner ID : 0000-5e01-1103  
Priority : 85  
Active : FE80::2

Forwarder 03

State : Active  
Virtual MAC : 000f-e2ff-4013 (Owner)  
Owner ID : 0000-5e01-1105  
Priority : 255  
Active : local

Forwarder Weight Track Information:

Track Object : 1 State : Positive Weight Reduced : 250

The output shows that when VLAN interface 3 on Switch A fails, the weight of the AVF on Switch A decreases to 5, which is lower than the lower limit of failure. All VFs on Switch A turn to initialized state and cannot be used for packet forwarding. The VF corresponding to MAC address 000f-e2ff-4011 on Switch C becomes the AVF, and Switch C takes over Switch A for packet forwarding.

# When the timeout timer (about 1800 seconds) expires, display the detailed information about VRRP group 1 on Switch C.

```
[SwitchC-Vlan-interface2] display vrrp ipv6 verbose
```

IPv6 Standby Information:

Run Mode : Load Balance  
Run Method : Virtual MAC

Total number of virtual routers : 1

Interface Vlan-interface2

VRID : 1 Adver Timer : 100  
Admin Status : Up State : Backup  
Config Pri : 100 Running Pri : 100  
Preempt Mode : Yes Delay Time : 5  
Become Master : 4200ms left  
Auth Type : None  
Virtual IP : FE80::10  
1::10  
Member IP List : FE80::3 (Local, Backup)  
FE80::1 (Master)  
FE80::2 (Backup)

Forwarder Information: 2 Forwarders 1 Active

Config Weight : 255  
Running Weight : 255

Forwarder 02

State : Listening  
Virtual MAC : 000f-e2ff-4012 (Learnt)  
Owner ID : 0000-5e01-1103  
Priority : 127  
Active : FE80::2

Forwarder 03



```

State : Active
Virtual MAC : 000f-e2ff-4013 (Owner)
Owner ID : 0000-5e01-1105
Priority : 255
Active : local
Forwarder Weight Track Information:
Track Object : 1 State : Positive Weight Reduced : 250

```

The output shows that when the timeout timer expires, the VF corresponding to virtual MAC address 000f-e2ff-4011 is removed, and does not forward the packets destined for the MAC address any more.

# When Switch A fails, use the **display vrrp ipv6 verbose** command to display the detailed information about VRRP group 1 on Switch B.

```
[SwitchB-Vlan-interface2] display vrrp ipv6 verbose
```

```

IPv6 Standby Information:
Run Mode : Load Balance
Run Method : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
VRID : 1 Adver Timer : 100
Admin Status : Up State : Master
Config Pri : 110 Running Pri : 110
Preempt Mode : Yes Delay Time : 5
Auth Type : None
Virtual IP : FE80::10
 1::10
Member IP List : FE80::2 (Local, Master)
 FE80::3 (Backup)
Forwarder Information: 2 Forwarders 1 Active
Config Weight : 255
Running Weight : 255
Forwarder 02
State : Active
Virtual MAC : 000f-e2ff-4012 (Owner)
Owner ID : 0000-5e01-1103
Priority : 255
Active : local
Forwarder 03
State : Listening
Virtual MAC : 000f-e2ff-4013 (Learnt)
Owner ID : 0000-5e01-1105
Priority : 127
Active : FE80::3
Forwarder Weight Track Information:
Track Object : 1 State : Positive Weight Reduced : 250

```

The output shows that when Switch A fails, Switch B becomes the master because its priority is higher than that of Switch C.

# Troubleshooting VRRP

## The screen frequently displays error prompts.

### Analysis

This error is probably caused by:

- Inconsistent configuration of the devices in the VRRP group.
- A device is attempting to send illegitimate VRRP packets.

### Solution

- In the first case, modify the configuration.
- In the latter case, resort to non-technical measures.

## Multiple masters are present in the same VRRP group.

### Analysis

- Multiple masters coexist for a short period: This is normal and requires no manual intervention.
- Multiple masters coexist for a long period: This is because devices in the VRRP group cannot receive VRRP packets, or the received VRRP packets are illegal.

### Solution

Ping between these masters, and do the following:

- If the ping fails, check network connectivity.
- If the ping succeeds, check that their configurations are consistent in terms of number of virtual IP addresses, virtual IP addresses, advertisement interval, and authentication.

## Frequent VRRP state transition.

### Analysis

The VRRP advertisement interval is set too short.

### Solution

Increase the interval to sent VRRP advertisement or introduce a preemption delay.

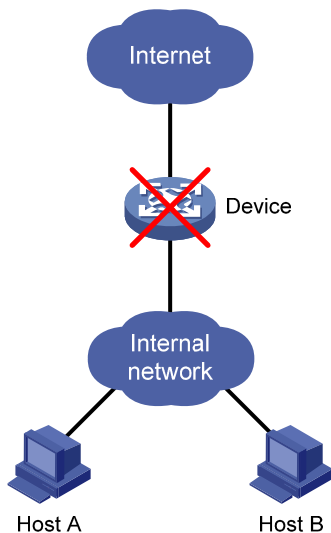
---

# Configuring stateful failover (available only on the HP 5500 EI)

## Stateful failover overview

Some customers require the key entries or access points of their networks, such as the Internet access point of an enterprise or a database server of a bank, to be highly reliable to ensure continuous data transmission. Deploying only one device (even with high reliability) in such a network risks a single point of failure, as shown in [Figure 48](#). Stateful failover can solve this problem.

**Figure 48 Network with one device deployed**



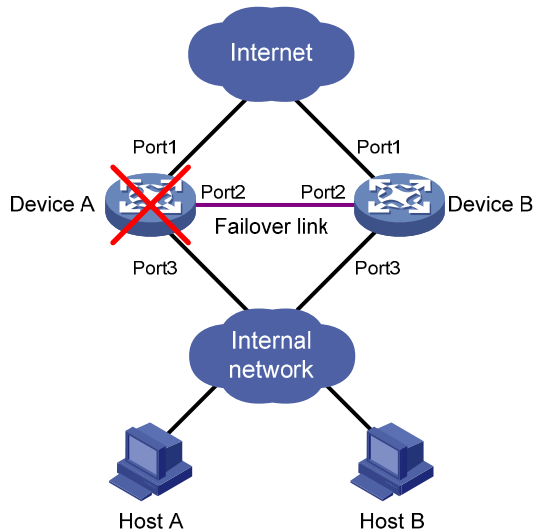
## Operating procedure

Stateful failover involves service backup and traffic switchover. The operating procedure of stateful failover is as follows:

1. As shown in [Figure 49](#), Device A and Device B connects to each other over a failover link.
2. The two devices exchange state negotiation messages periodically through the failover link. After the two devices enter the synchronized state, they back up the sessions of each other to make sure that the sessions on them are consistent.
3. If one device fails, the other device can take over the services by using VRRP or a dynamic routing protocol (such as OSPF) to avoid service interruption.

In this document, the stateful failover feature supports backing up only portal services.

Figure 49 Network diagram for stateful failover

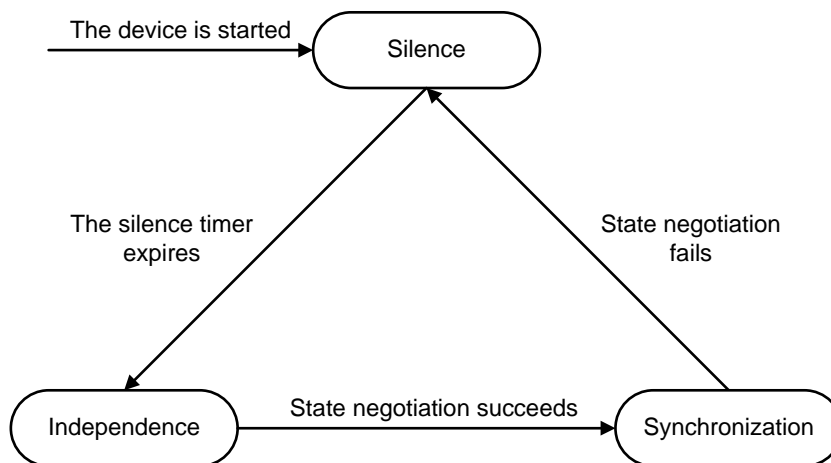


## Stateful failover states

Stateful failover has the following states:

- **Silence**—Indicates that the device has just started, or is transiting from synchronization state to independence state.
- **Independence**—Indicates that the silence timer has expired, but no failover link is established.
- **Synchronization**—Indicates that the device has completed state negotiation with the other device and is ready for service backup.

Figure 50 Stateful failover state relations



## Introduction to stateful failover configuration

To implement stateful failover on two devices, perform the following configurations:

- **Routing configuration.** Configure VRRP or a dynamic routing protocol on the devices and the uplink/downlink devices to make sure that the traffic can automatically switch to the other device when a device fails.

- Service backup configuration, which can implement real-time service backup between the two devices.

This configuration guide only introduces the service backup configuration.

Complete the following tasks to configure stateful failover:

Task	Remarks
<a href="#">Enabling stateful failover</a>	Required
<a href="#">Configuring the backup VLAN</a>	Required
Service module related configurations	Optional You must perform further configurations on the device before it can automatically back up portal service information to the backup device. For more information, see <i>Security Configuration Guide</i> .

## Enabling stateful failover

When you enable stateful failover with the **dhbk enable backup-type { dissymmetric-path | symmetric-path }** command:

- If you specify the **dissymmetric-path** keyword, the two devices operate in active/active mode. Sessions enter and leave the internal network through different devices to achieve load sharing.
- If you specify the **symmetric-path** keyword, the two devices operate in active/standby mode. Sessions enter and leave the internal network through one device.

Select a keyword based on the network environment and resources, and specify the same keyword for both devices.

To enable stateful failover:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Enable stateful failover in a specified mode.	<b>dhbk enable backup-type { dissymmetric-path   symmetric-path }</b>	Disabled by default.

## Configuring the backup VLAN

After you specify a VLAN as a backup VLAN, the interfaces added to the VLAN can serve as stateful failover interfaces to transmit stateful failover packets.

The device identifies stateful failover packets by the VLAN tag and private protocol number, and broadcasts them in the backup VLAN to the peer. Do not configure other services for the backup VLAN (such as MAC VLAN or Voice VLAN); otherwise, the operation of stateful failover may be affected.

The interfaces assigned to a backup VLAN can forward other packets besides stateful failover packets.

To configure a backup VLAN:

Step	Command	Remarks
1. Enter system view.	<b>system-view</b>	N/A
2. Create a VLAN and assign interfaces to the VLAN.	See <i>Layer 2—LAN Switching Configuration Guide</i> .	N/A
3. Return to system view.	<b>quit</b>	N/A
4. Specify the VLAN as a backup VLAN.	<b>dhbk vlan <i>vlan-id</i></b>	Not specified by default.

## Displaying and maintaining stateful failover

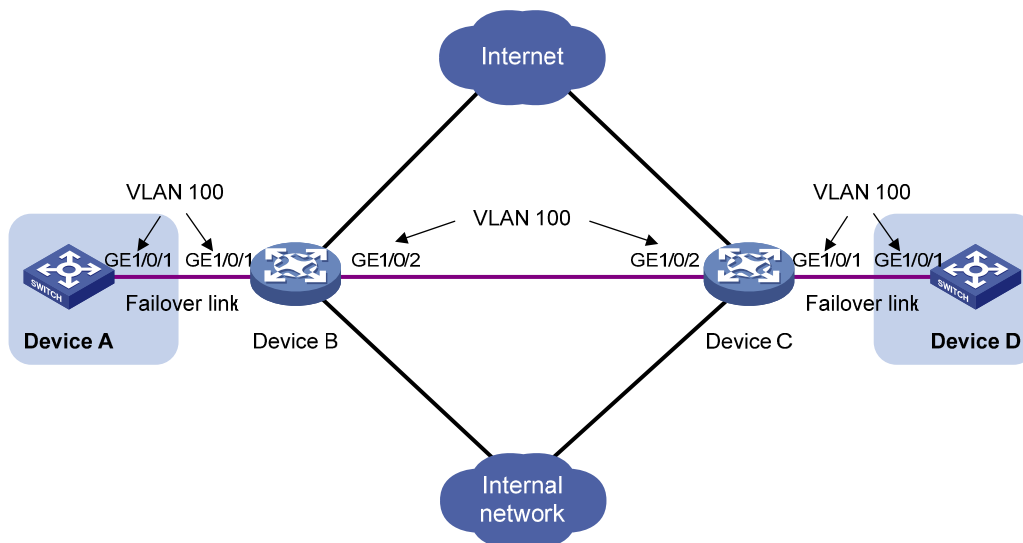
Task	Command	Remarks
Display the running status and related information of stateful failover.	<b>display dhbk status</b> [   { <b>begin</b>   <b>exclude</b>   <b>include</b> } <i>regular-expression</i> ]	Available in any view

## Stateful failover configuration example

### Network requirements

In [Figure 51](#), Device B and Device C serve as the internal gateways of an enterprise network. Device A and Device D, respectively attached to Device B and Device C, provide portal access authentication for internal users. Configure stateful failover between Device A and Device D. When one device fails, the other device takes over the services to ensure service continuity.

**Figure 51 Network diagram**



### Configuration procedure

1. Configure Device A:  

```
Create VLAN 100.
<DeviceA> system-view
```

- ```
[DeviceA] vlan 100
# Assign GigabitEthernet 1/0/1 to VLAN 100.
[DeviceA-vlan100] port gigabitethernet 1/0/1
[DeviceA-vlan100] quit
# Specify VLAN 100 as a backup VLAN.
[DeviceA] dhbk vlan 100
# Enable symmetric-path mode stateful failover.
[DeviceA] dhbk enable backup-type symmetric-path
```
2. Configure Device B:

```
# Create VLAN 100.
<DeviceB> system-view
[DeviceB] vlan 100
# Assign GigabitEthernet 1/0/1 to VLAN 100.
[DeviceB-vlan100] port gigabitethernet 1/0/1
[DeviceB-vlan100] quit
# Assign GigabitEthernet 1/0/2 to VLAN 100.
Because Device B and Device C may exchange packets of multiple VLANs, configure
GigabitEthernet 1/0/2 as a trunk port and permit packets of VLAN 100 to pass.
[DeviceB] interface gigabitethernet 1/0/2
[DeviceB-GigabitEthernet1/0/2] port link-type trunk
[DeviceB-GigabitEthernet1/0/2] port trunk permit vlan 100
Please wait... Done.
```
 3. The configurations on Device C are similar to those on Device B. (Details not shown.)
 4. The configurations on Device D are similar to those on Device A. (Details not shown.)

Configuration guidelines

When you configure stateful failover, follow these guidelines:

- Stateful failover can be implemented only between two devices rather than among more than two devices.
- The same numbered interfaces must exist on the two devices. Otherwise, session backup fails. For example, if Device A uses GigabitEthernet 1/0/1 and GigabitEthernet 1/0/3 to forward backup data, Device B must also use GigabitEthernet 1/0/1 and GigabitEthernet 1/0/3.

Configuring BFD (available only on the HP 5500 EI)

The term *router* in the BFD features refers to both routers and Layer 3 switches.

The term *interface* in the BFD features refers to Layer 3 interfaces, including VLAN interfaces and route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).

BFD overview

Devices must quickly detect communication failures so that measures can be taken promptly to ensure service continuity and enhance network availability.

The main fault detection methods include the following:

- **Hardware detection**—Detects link failures by sending hardware detection signals, such as synchronous digital hierarchy (SDH) alarms. Hardware detection can quickly detect link failures, but not all media types support hardware detection.
- **Hello mechanism**—Devices can use the hello mechanism of a routing protocol to detect link failures, which has a failure detection rate in seconds. On a high-speed interface, such as a Gigabit interface, a failure that lasts for one second will cause a large quantity of data to be dropped. The hello mechanism is unacceptable for delay-sensitive services such as voice service. Moreover, this detection method largely relies on the routing protocol.
- **Other detection methods**—Some protocols provide dedicated detection mechanisms. However, they cannot be deployed for inter-system communications.

Bidirectional forwarding detection (BFD) provides a single mechanism to monitor links. With BFD, devices can quickly detect communication failures and restore communications through backup paths.

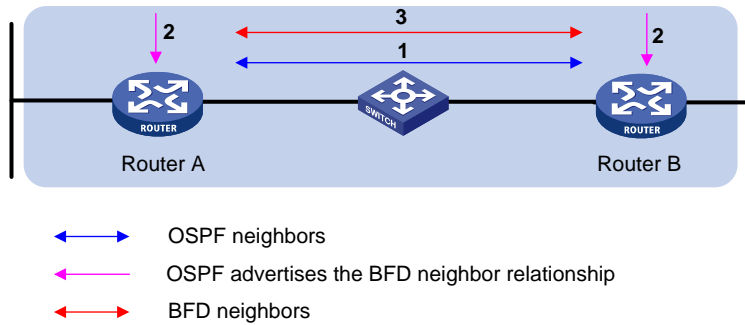
How BFD works

BFD provides a general-purpose, standard, medium- and protocol-independent fast failure detection mechanism. It can uniformly and quickly detect the failures of the bidirectional forwarding paths between two routers for protocols, such as routing protocols.

BFD provides no neighbor discovery mechanism. Protocols that BFD services notify BFD of routers to which it needs to establish sessions. After a session is established, if no BFD control packet is received from the peer within the negotiated BFD interval, BFD notifies a failure to the protocol, which takes appropriate measures.

Operation of BFD

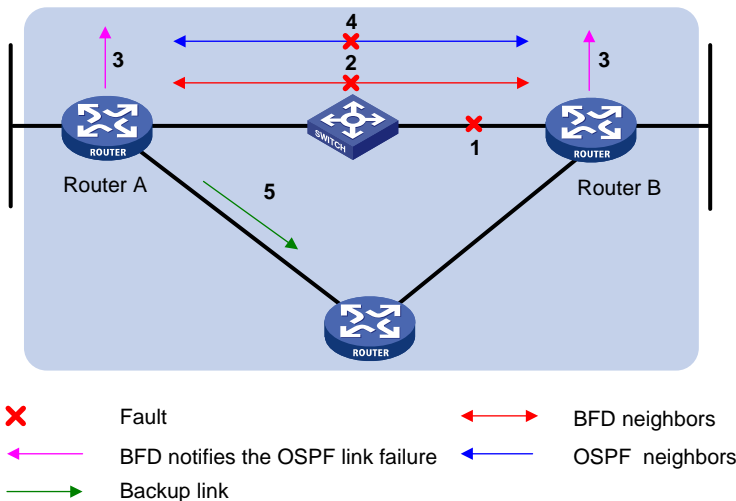
Figure 52 BFD session establishment (on OSPF routers)



The process of BFD session establishment is as follows:

1. A protocol sends hello messages to discover neighbors and establish neighborships.
2. After establishing neighborships, the protocol notifies BFD of the neighbor information, including destination and source addresses.
3. BFD uses the information to establish BFD sessions.

Figure 53 BFD fault detection (on OSPF routers)



The process of BFD fault detection is as follows:

1. BFD detects a link failure.
2. BFD clears the neighbor session.
3. BFD notifies the protocol of the failure.
4. The protocol terminates the neighborship on the link.
5. If a backup link is available, the protocol will use it to forward packets.

NOTE:

No detection time resolution is defined in the BFD draft. Most devices supporting BFD provide detection measured in milliseconds.

BFD detection methods

- **Single-hop detection**—Detects the IP connectivity between two directly connected systems.
- **Multi-hop detection**—Detects any of the paths between two systems. These paths have multiple hops and may be overlapped.
- **Bidirectional detection**—Sends detection packets at two sides of a bidirectional link to detect the bidirectional link status, finding link failures in milliseconds. (BFD LSP detection is a special case in which BFD control packets are sent in one direction, and the peer device reports the link status through other links.)

BFD session modes

- **Control packet mode**—Both ends of the link exchange BFD control packets to monitor link status.
- **Echo mode**—One end of the link sends Echo packets to the other end, which then forwards the packets back to the originating end, monitoring link status in both directions.

BFD operating modes

Before a BFD session is established, BFD has the following operating modes—active and passive.

- **Active mode**—BFD actively sends BFD control packets regardless of whether any BFD control packet is received from the peer.
- **Passive mode**—BFD does not send control packets until a BFD control packet is received from the peer.

At least one end must operate in the active mode for a BFD session to be established.

After a BFD session is established, both ends must operate in the asynchronous mode—Both endpoints periodically send BFD control packets to each other. BFD considers the session down if it receives no BFD control packets within a specific interval.

NOTE:

When a BFD session is maintained by sending Echo packets, the session is independent of the operating mode.

Dynamic BFD parameter changes

After a BFD session is established, both ends can negotiate the related BFD parameters, such as the minimum transmit interval, minimum receive interval, initialization mode, and packet authentication mode. After that, both ends use the negotiated parameters, without affecting the current session state.

Authentication modes

BFD provides the following authentication methods:

- **Simple**—Simple authentication
- **MD5**—MD5 (Message Digest 5) authentication
- **SHA1**—SHA1 (Secure Hash Algorithm 1) authentication

BFD packet format

BFD control packets are encapsulated into UDP packets with port number 3784 for single-hop detection or port number 4784 for multi-hop detection (it can also be 3784 based on the configuration task). BFD echo packets have a similar format to BFD control packets (except that the *Desired Min TX Interval* and *Required Min RX Interval* fields are null), with UDP port number 3785.

Figure 54 BFD packet format

| | | | |
|-------------------------------|------|----------|------------------------|
| 0 | 7 | 23 | 31 |
| Vers | Diag | Sta | P F C A D R |
| Detect Mult | | Length | |
| My Discriminator | | | |
| Your Discriminator | | | |
| Desired Min TX Interval | | | |
| Required Min RX Interval | | | |
| Required Min Echo RX Interval | | | |
| Auth Type | | Auth Len | Authentication Data... |

- **Vers**—Protocol version. The protocol version is 1.
- **Diag**—This bit indicates the reason for the last transition of the local session from **up** to some other state. Table 19 lists the states.

Table 19 Diag bit values

| Diag | Description |
|------|--------------------------------|
| 0 | No Diagnostic |
| 1 | Control Detection Time Expired |
| 2 | Echo Function Failed |
| 3 | Neighbor Signaled Session Down |
| 4 | Forwarding Plane Reset |
| 5 | Path Down |
| 6 | Concatenated Path Down |
| 7 | Administratively Down |
| 8 | Reverse Concatenated Path Down |
| 9~31 | Reserved for future use |

- **State (Sta)**—Current BFD session state. Its value can be 0 for AdminDown, 1 for Down, 2 for Init, and 3 for Up.
- **Poll (P)**—If set, the transmitting system is requesting verification of connectivity, or of a parameter change. If clear, the transmitting system is not requesting verification.
- **Final (F)**—If set, the transmitting system is responding to a received BFD control packet that had the Poll (P) bit set. If clear, the transmitting system is not responding to a Poll.
- **Control Plane Independent (C)**—If set, the transmitting system's BFD implementation does not share fate with its control plane (BFD is implemented in the forwarding plane and can continue to function through disruptions in the control plane.) If clear, the transmitting system's BFD implementation shares fate with its control plane.
- **Authentication Present (A)**—If set, the Authentication Section is present, and the session is to be authenticated.
- **Demand (D)**—If set, Demand mode is active in the transmitting system (the system wishes to operate in Demand mode, knows that the session is up in both directions, and is directing the remote system to cease the periodic transmission of BFD Control packets). If clear, Demand mode is not active in the transmitting system.

- **Reserved (R)**—This byte must be set to zero on transmit and ignored on receipt.
- **Detect Mult**—Detection time multiplier.
- **Length**—Length of the BFD control packet, in bytes.
- **My Discriminator**—A unique, nonzero discriminator value generated by the transmitting system, used to demultiplex multiple BFD sessions between the same pair of systems.
- **Your Discriminator**—The discriminator received from the remote system. This field reflects back the received value of My Discriminator or is 0 if that value is unknown.
- **Desired Min TX Interval**—This is the minimum interval, in microseconds, that the local system would like to use when transmitting BFD control packets. The value zero is reserved.
- **Required Min RX Interval**—This is the minimum interval, in microseconds, between received BFD control packets that this system is capable of supporting. If this value is zero, the transmitting system does not want the remote system to send any periodic BFD control packets.
- **Required Min Echo RX Interval**—This is the minimum interval, in microseconds, between received BFD echo packets that this system is capable of supporting. If this value is zero, the transmitting system does not support the receipt of BFD echo packets.
- **Auth Type**—The authentication type in use, if the Authentication Present (A) bit is set.
- **Auth Len**—The length, in bytes, of the authentication section, including the Auth Type and Auth Len fields.

Supported features

- OSPF. For more information, see *Layer 3—IP Routing Configuration Guide*.
- OSPFv3. For more information, see *Layer 3—IP Routing Configuration Guide*.
- IS-IS. For more information, see *Layer 3—IP Routing Configuration Guide*.
- IPv6 IS-IS. For more information, see *Layer 3—IP Routing Configuration Guide*.
- RIP. For more information, see *Layer 3—IP Routing Configuration Guide*.
- Static routing. For more information, see *Layer 3—IP Routing Configuration Guide*.
- BGP. For more information, see *Layer 3—IP Routing Configuration Guide*.
- IPv6 BGP. For more information, see *Layer 3—IP Routing Configuration Guide*.
- PIM. For more information, see *IP Multicast Configuration Guide*.
- IPv6 PIM. For more information, see *IP Multicast Configuration Guide*.
- Track. For more information, see "[Configuring track](#)."
- IP fast reroute (FRR). Currently, IP FRR is supported by OSPF, RIP, IS-IS, and static routing. For more information, see *Layer 3—IP Routing Configuration Guide*.

Protocols and standards

- draft-ietf-bfd-base-09, *Protocol Independent Bidirectional Forwarding Detection*
- draft-ietf-bfd-v4v6-1 hop-10, *BFD for IPv4 and IPv6 (Single Hop)*
- draft-ietf-bfd-multihop-08, *BFD for Multihop Paths*
- draft-ietf-bfd-generic-05, *Generic Application of BFD*

Configuring BFD basic functions

The BFD basic functions configuration is the basis for configuring BFD for other protocols.

Configuration prerequisites

Before configuring BFD basic functions, complete the following tasks:

- Configure the network layer addresses of the interfaces so that adjacent nodes are reachable to each other at the network layer
- Configure the routing protocols that support BFD

Configuration procedure

To configure BFD basic functions:

| Step | Command | Remarks |
|---|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Specify the mode for establishing a BFD session. | bfd session init-mode { active passive } | Optional.
active by default. |
| 3. Configure the destination port number for multi-hop BFD control packets. | bfd multi-hop destination-port port-number | Optional.
4784 by default. |
| 4. Configure the source IP address of echo packets. | bfd echo-source-ip ip-address | Optional.
The source IP address should not be on the same network segment as any local interface's IP address. Otherwise, a large number of ICMP redirect packets may be sent from the peer, resulting in link congestion. |
| 5. Enter interface view. | interface interface-type interface-number | N/A |
| 6. Configure the minimum interval for receiving BFD echo packets. | bfd min-echo-receive-interval value | Optional.
For relevant information, see the description of the <i>Required Min Echo RX Interval</i> field in " BFD packet format ."
400 milliseconds by default. |
| 7. Configure the minimum interval for transmitting BFD control packets. | bfd min-transmit-interval value | Optional.
For relevant information, see the description of the <i>Desired Min TX Interval</i> field in " BFD packet format ."
400 milliseconds by default. |

| Step | Command | Remarks |
|--|---|---|
| 8. Configure the minimum interval for receiving BFD control packets. | bfd min-recv-interval <i>value</i> | Optional.
For relevant information, see the description of the <i>Required Min RX Interval</i> field in "BFD packet format."
400 milliseconds by default. |
| 9. Configure the detection time multiplier. | bfd detect-multiplier <i>value</i> | Optional.
For relevant information, see the description of the <i>Detect Mult</i> field in "BFD packet format."
5 by default. |
| 10. Configure the authentication mode on the interface. | bfd authentication-mode { md5 <i>key-id</i> [cipher] <i>key</i> sha1 <i>key-id</i> [cipher] <i>key</i> simple <i>key-id</i> [cipher] <i>password</i> } | Optional.
By default, the interface operates in the non-authentication mode. |

In Figure 52 for example, if you configure the Desired Min TX Interval as 100 milliseconds, Required Min RX Interval as 300 milliseconds, and Detect Mult as 5 on Router A, and configure the Desired Min TX Interval as 150 milliseconds, Required Min RX Interval as 400 milliseconds, and Detect Mult as 10 on Router B,

- The actual transmitting interval on Router A is 400 milliseconds, which is the greater value between the minimum interval for transmitting BFD control packets on Router A (100 milliseconds) and the minimum interval for receiving BFD control packets on Router B (400 milliseconds).
- The actual transmitting interval on Router B is 300 milliseconds, which is the greater value between the minimum interval for transmitting BFD control packets on Router B (150 milliseconds) and the minimum interval for receiving BFD control packets on Router A (300 milliseconds).
- The actual detection time on Router A is 3000 milliseconds, which is 10×300 milliseconds (Detect Mult on Router B \times actual transmitting interval on Router B).
- The actual detection time on Router B is 2000 milliseconds, which is 5×400 milliseconds (Detect Mult on Router A \times actual transmitting interval on Router A).

Displaying and maintaining BFD

| Task | Command | Remarks |
|---|---|------------------------|
| Display information about BFD-enabled interfaces. | display bfd interface [verbose] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display information about enabled BFD debugging. | display bfd debugging-switches [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display BFD session information. | display bfd session [slot <i>slot-number</i> [all verbose] verbose] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Clear BFD session statistics. | reset bfd session statistics [slot <i>slot-number</i>] | Available in user view |

Configuring track

Only the HP 5500 EI Switch Series supports BFD, VRRP, and PBR configurations.

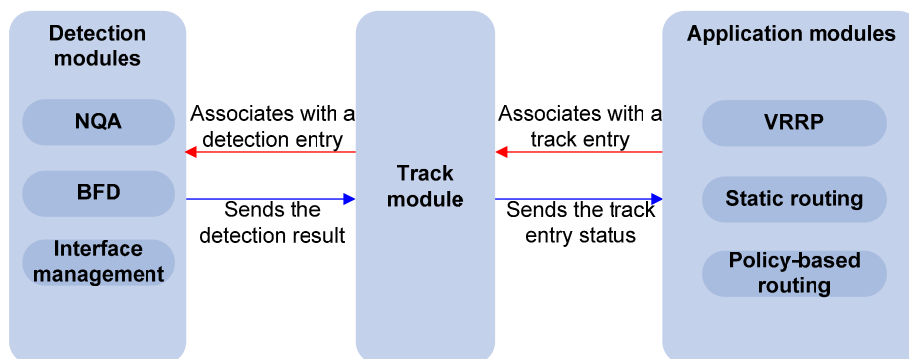
Track overview

Introduction to collaboration

The track module works between application and detection modules, as shown in [Figure 55](#). It shields the differences between various detection modules from application modules.

Collaboration is enabled after you associate the track module with a detection module and an application module. The detection module probes specific objects such as interface status, link status, network reachability, and network performance, and informs the track module of detection results. The track module sends the detection results to the associated application module. When notified of changes of the tracked object, the application modules can react to avoid communication interruption and network performance degradation.

Figure 55 Collaboration through the track module



Collaboration fundamentals

The track module collaborates with detection modules and application modules:

- Collaboration between the track module and a detection module
- Collaboration between the track module and an application module

Collaboration between the track module and a detection module

The detection module sends the detection result of the associated tracked object to the track module. Depending on the result, the track module changes the status of the track entry:

- If the tracked object functions normally, for example, the target interface is up or the target network is reachable, the state of the track entry is Positive.
- If the tracked object functions abnormally, for example, the target interface is down or the target network is unreachable, the state of the track entry is Negative.
- If the detection result is not valid, for example, the NQA test group that is associated with the track entry does not exist, the state of the track entry is Invalid.

The following detection modules can be associated with the track module:

- NQA
- BFD (available only on the HP 5500 EI)
- Interface management module

Collaboration between the track module and an application module

After being associated with an application module, when the status of the track entry changes, the track module notifies the application module, which then takes proper actions.

The following application modules can be associated with the track module:

- Virtual Router Redundancy Protocol (VRRP) (available only on the HP 5500 EI)
- Static routing
- Policy-based routing (available only on the HP 5500 EI)

In some cases, the status of a track entry changes while a route is still recovering. This leads to problems if the track module immediately notifies the application modules of the status change and the application modules begin using the route before it is ready.

For example, the master in a VRRP group monitors the uplink interface through the track module. When the uplink interface fails, the track module notifies the master to reduce its priority so that a backup with a higher priority can preempt as the master to forward packets. When failed uplink interface recovers, if the track module immediately notifies the original master to restore its priority, the master immediately will forward packets to that interface; however, this result in packet forwarding failure because the uplink route has not yet been recovered.

To solve this problem, configure a delay before the track module notifies the application modules of the track entry status changes.

Collaboration application example

The following is an example of collaboration between NQA, track, and static routing.

Configure a static route with next hop 192.168.0.88 on the device. If the next hop is reachable, the static route is valid. If the next hop becomes unreachable, the static route should become invalid. For this purpose, configure collaboration between the NQA, track, and static routing modules:

1. Create an NQA test group to monitor the reachability of IP address 192.168.0.88.
2. Create a track entry and associate it with the NQA test group. When the next hop 192.168.0.88 is reachable, the track entry is in Positive state. When the next hop becomes unreachable, the track entry is in Negative state.
3. Associate the track entry with the static route. When the track entry turns to the Positive state, the static route is valid. When the associated track entry turns to the Negative state, the static route is invalid.

Track configuration task list

To implement the collaboration function, establish associations between the track module and the detection modules, and between the track module and the application modules.

Complete these tasks to configure the track module:

| Task | Remarks |
|---|--|
| Associating the track module with a detection module | Associating track with NQA |
| | Associating track with BFD (available only on the HP 5500 EI) |
| | Associating track with interface management |
| Associating the track module with an application module | Associating track with VRRP (available only on the HP 5500 EI) |
| | Associating track with static routing |
| | Associating track with PBR (available only on the HP 5500 EI) |

Associating the track module with a detection module

Associating track with NQA

NQA supports multiple test types to analyze network performance, services, service quality. For example, an NQA test group can periodically detect whether a destination is reachable, or whether the TCP connection to a TCP server can be set up.

An NQA test group functions as follows when it is associated with a track entry:

- If the consecutive failures reach the specified threshold, the NQA module tells the track module that the tracked object malfunctions. Then the track module sets the track entry to the Negative state.
- If the specified threshold is not reached, the NQA module tells the track module that the tracked object functions normally. The track module then sets the track entry to the Positive state.

For more information about NQA, see *Network Management and Monitoring Configuration Guide*.

To associate track with NQA:

| Step | Command | Remarks |
|--|---|---------------------------------------|
| 1. Enter system view. | system-view | N/A |
| 2. Create a track entry, associate it with an NQA reaction entry, and specify the delay time for the track module to notify the associated application module when the track entry status changes. | track track-entry-number nqa entry admin-name operation-tag reaction item-number [delay { negative negative-time positive positive-time } *] | No track entry is created by default. |

NOTE:

If the specified NQA test group or the reaction entry in the track entry does not exist, the status of the track entry is Invalid.

Associating track with BFD (available only on the HP 5500 EI)

BFD supports the control packet mode and echo mode. Only echo-mode BFD can be associated with a track entry.

The BFD functions as follows when it is associated with a track entry:

- If the BFD detects that the link fails, it informs the track entry of the link failure. The track module then sets the track entry to the Negative state.
- If the BFD detects that the link is normal, the track module sets the track entry to the Positive state.

For more information about BFD, see "[Configuring BFD \(available only on the HP 5500 EI\)](#)."

Configuration prerequisites

Before you associate track with BFD, configure the source address of the BFD echo packets.

Configuration procedure

To associate track with BFD:

| Step | Command | Remarks |
|--|--|---------------------------------------|
| 1. Enter system view. | system-view | N/A |
| 2. Create a track entry, associate it with the BFD session, and specify the delay time for the track module to notify the associated application module when the track entry status changes. | track track-entry-number bfd echo interface interface-type interface-number remote ip remote-ip local ip local-ip [delay { negative negative-time positive positive-time } *] | No track entry is created by default. |

NOTE:

When associating track with BFD, do not configure the virtual IP address of a VRRP group as the local or remote address of a BFD session.

Associating track with interface management

The interface management module monitors the physical status or network-layer protocol status of the interface. The interface management module functions as follows when it is associated with a track entry:

- When the physical or network-layer protocol status of the interface changes to up, the interface management module informs the track module of the change and the track module sets the track entry to Positive.
- When the physical or network-layer protocol status of the interface changes to down, the interface management module informs the track module of the change and the track module sets the track entry to Negative.

To associate track with interface management:

| Step | Command | Remarks |
|-----------------------|--------------------|---------|
| 1. Enter system view. | system-view | N/A |

| Step | Command | Remarks |
|---|---|--|
| 2. Associate track with interface management. | <p>Create a track entry, associate it with the interface management module to monitor the physical status of an interface, and specify the delay time for the track module to notify the associated application module when the track entry status changes:</p> <pre>track track-entry-number interface interface-type interface-number [delay { negative negative-time positive positive-time } *]</pre> <p>Create a track entry, associate it with the interface management module to monitor the Layer 3 protocol status of an interface, and specify the delay time for the track module to notify the associated application module when the track entry status changes:</p> <pre>track track-entry-number interface interface-type interface-number protocol { ipv4 ipv6 } [delay { negative negative-time positive positive-time } *]</pre> | <p>Use either approach.</p> <p>No track entry is created by default.</p> |

Associating the track module with an application module

Associating track with VRRP (available only on the HP 5500 EI)

VRRP is an error-tolerant protocol. It adds a group of routers that can act as network gateways to a VRRP group, which forms a virtual router. Routers in the VRRP group elect the master acting as the gateway according to their priorities. A router with a higher priority is more likely to become the master. The other routers function as the backups. When the master fails, the backups in the VRRP group elect a new gateway to undertake the responsibility of the failed master. This ensures that the hosts in the network segment can uninterruptedly communicate with external networks.

When VRRP is operating in standard protocol mode or load balancing mode, associate the track module with the VRRP group to implement the following actions:

- Change the priority of a router according to the status of the uplink. If a fault occurs on the uplink of the router, the VRRP group cannot be aware of the uplink failure. If the router is the master, hosts in the LAN cannot access the external network. This problem can be solved by establishing a track-VRRP group association. Use the detection modules to monitor the status of the uplink of the router and establish collaborations between the detection modules, track module and VRRP. When the uplink fails, the detection modules notify the track module to change the status of the monitored track entry to Negative, and the priority of the master then decreases by a specific value, allowing a higher priority router in the VRRP group to become the master, and maintaining proper communication between the hosts in the LAN and the external network.
- Monitor the master on a backup. If a fault occurs on the master, the backup operating in switchover mode will switch to the master immediately to maintain normal communication.

When VRRP is operating in load balancing mode, associate the track module with the VRRP Virtual Forwarder (VF) to implement the following functions:

- Change the priority of the active VF (AVF) according to its uplink state. When the uplink of the AVF fails, the track entry changes to Negative state and the weight of the AVF decreases by a specific value so that the VF with a higher priority becomes the new AVF to forward packets.

- Monitor the AVF status from the listening VF (LVF), which refers to the VF in listening state. When the AVF fails, the LVF that is operating in switchover mode becomes the new AVF to ensure continuous forwarding.

VRRP tracking is not valid on an IP address owner. An IP address owner refers to a router when the IP address of the virtual router is the IP address of an interface on the router in the VRRP group.

For more information about VRRP, see "[Configuring VRRP \(available only on the HP 5500 EI\)](#)."

To associate track with VRRP group:

| Step | Command | Remarks |
|--|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter interface view. | interface <i>interface-type</i>
<i>interface-number</i> | N/A |
| 3. Create a VRRP group and configure its virtual IP address. | vrrp vrid <i>virtual-router-id</i> virtual-ip
<i>virtual-address</i> | No VRRP group is created by default. |
| 4. Associate a track entry with a VRRP group. | vrrp [ipv6] vrid <i>virtual-router-id</i>
track <i>track-entry-number</i> [reduced
<i>priority-reduced</i> switchover] | No track entry is specified for a VRRP group by default.
This command is supported when VRRP is operating in both standard protocol mode and load balancing mode. |

To associate track with VRRP VF:

| Step | Command | Remarks |
|--|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter interface view. | interface <i>interface-type</i>
<i>interface-number</i> | N/A |
| 3. Create a VRRP group and configure its virtual IP address. | vrrp vrid <i>virtual-router-id</i> virtual-ip
<i>virtual-address</i> | No VRRP group is created by default. |
| 4. Associate track with VRRP VF. | Associate a track entry with the VRRP VF:
vrrp [ipv6] vrid <i>virtual-router-id</i>
weight track <i>track-entry-number</i>
[reduced <i>weight-reduced</i>]
Configure the LVF to monitor the AVF status through the track entry:
vrrp [ipv6] vrid <i>virtual-router-id</i>
track <i>track-entry-number</i>
forwarder-switchover member-ip
<i>ip-address</i> | Use at least one command.
By default, no track entry is specified for a VF.
This command is configurable when VRRP is operating in standard mode or load balancing mode. However, this function takes effect only when VRRP is operating in load balancing mode. |

NOTE:

- When the status of the track entry changes from Negative to Positive or Invalid, the associated router or VF restores its priority automatically.
- You can associate a nonexistent track entry with a VRRP group or VF. The association takes effect only after you use the **track** command to create the track entry.

Associating track with static routing

A static route is a manually configured route. With a static route configured, packets to the specified destination are forwarded through the path specified by the administrator.

The disadvantage of using static routes is that they cannot adapt to network topology changes. Faults or topological changes in the network can make the routes unreachable, causing network breaks.

To prevent this problem, configure another route to back up the static route. When the static route is reachable, packets are forwarded through the static route. When the static route is unreachable, packets are forwarded through the backup route, avoiding network breaks and enhancing network reliability.

To check the accessibility of a static route in real time, establish association between the track and the static route.

If you specify the next hop but not the egress interface when configuring a static route, you can establish collaborations among the static route, the track module, and detection modules. This enables you to check the accessibility of the static route by the status of the track entry.

- The Positive state of the track entry shows that the next hop of the static route is reachable and that the configured static route is valid.
- The Negative state of the track entry shows that the next hop of the static route is not reachable and that the configured static route is invalid.
- The Invalid state of the track entry shows that the accessibility of the next hop of the static route is unknown and that the static route is valid.

If the track module detects the next hop accessibility of the static route in a private network through NQA, the VPN instance name of the next hop of the static route must be consistent with that configured for the NQA test group. Otherwise, the accessibility detection cannot function properly.

If a static route needs route recursion, the associated track entry must monitor the next hop of the recursive route instead of that of the static route; otherwise, a valid route may be considered invalid.

For more information about static route configuration, see *Layer 3—IP Routing Configuration Guide*.

To associate track with static routing:

| Step | Command | Remarks |
|--|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Associate the static route with a track entry to check the accessibility of the next hop. | <p>Approach 1:</p> <pre>ip route-static dest-address { mask mask-length } { next-hop-address vpn-instance d-vpn-instance-name next-hop-address } track track-entry-number [preference preference-value] [tag tag-value] [description description-text]</pre> <p>Approach 2:</p> <pre>ip route-static vpn-instance s-vpn-instance-name&<1-6> dest-address { mask mask-length } { next-hop-address track track-entry-number [public] vpn-instance d-vpn-instance-name next-hop-address track track-entry-number } [preference preference-value] [tag tag-value] [description description-text]</pre> | <p>Use either approach.</p> <p>Not configured by default.</p> <p>Only the HP 5500 EI Switch Series supports the vpn-instance keyword.</p> |

NOTE:

You can associate a nonexistent track entry with a static route. The association takes effect only after you use the **track** command to create the track entry.

Associating track with PBR (available only on the HP 5500 EI)

Policy-based routing (PBR) is a routing mechanism based on user-defined policies. Different from the traditional destination-based routing mechanism, PBR enables you to use a policy (based on the source address and other criteria) to route packets.

PBR cannot detect the availability of any action taken on packets. When an action is not available, packets processed by the action may be discarded. For example, configure PBR to forward packets that match certain criteria through a specific next hop. When the specified next hop fails, PBR cannot sense the failure, and continues to forward matching packets to the next hop.

This problem can be solved by associating track with PBR, which improves the flexibility of PBR application, and enables PBR to sense topology changes.

After you associate a track entry with an apply clause, the detection module associated with the track entry sends the detection result of the availability of the object (an interface or an IP address) specified in the apply clause.

- The Positive state of the track entry shows that the object is available, and the apply clause is valid.
- The Negative state of the track entry shows that the object is not available, and the apply clause is invalid.
- The Invalid state of the track entry shows that the apply clause is valid.

The following objects can be associated with a track entry:

- Next hop
- Default next hop

For more information about PBR, see *Layer 3—IP Routing Configuration Guide*.

Configuration prerequisites

Before you associate track with PBR, create a policy or a policy node and configure the match criteria as well.

Configuration procedure

To associate track with PBR:

| Step | Command | Remarks |
|---|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Create a policy or policy node and enter PBR policy node view. | policy-based-route <i>policy-name</i> [deny permit] node <i>node-number</i> | Required |
| 3. Define an ACL match criterion. | if-match acl <i>acl-number</i> | Optional
By default, no packets are filtered. |

| Step | Command | Remarks |
|------------------------------|---|---|
| 4. Associate track with PBR. | Set the next hop and associate it with a track entry:
apply ip-address next-hop <i>ip-address</i>
[direct] [track <i>track-entry-number</i>]
[<i>ip-address</i> [direct] [track <i>track-entry-number</i>]]

Set the default next hop, and associate it with a track entry:
apply ip-address default next-hop
<i>ip-address</i> [track <i>track-entry-number</i>]
[<i>ip-address</i> [track <i>track-entry-number</i>]] | Configure at least one of the commands. |

NOTE:

You can associate a nonexistent track entry with PBR. The association takes effect only after you use the **track** command to create the track entry.

Displaying and maintaining track entries

| Task | Command | Remarks |
|---|--|-----------------------|
| Display information about the specified or all track entries. | display track { <i>track-entry-number</i> all } [[{ begin exclude include } <i>regular-expression</i>]] | Available in any view |

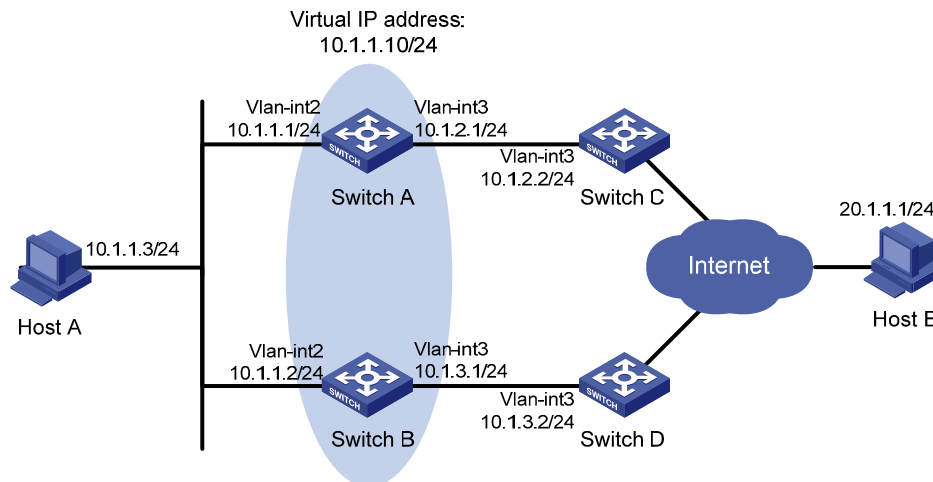
Track configuration examples

VRRP-track-NQA collaboration configuration example (the master monitors the uplink) (available only on the HP 5500 EI)

Network requirements

- As shown in [Figure 56](#), Host A needs to access Host B on the Internet. The default gateway of Host A is 10.1.1.10/24.
- Switch A and Switch B belong to VRRP group 1, whose virtual IP address is 10.1.1.10.
- When Switch A works properly, packets from Host A to Host B are forwarded through Switch A. When VRRP finds that a fault is on the uplink of Switch A through NQA, packets from Host A to Host B are forwarded through Switch B.

Figure 56 Network diagram



Configuration procedure

1. Create VLANs, and assign corresponding ports to the VLANs. Configure the IP address of each VLAN interface as shown in Figure 56. (Details not shown.)

2. Configure an NQA test group on Switch A:

```
<SwitchA> system-view
# Create an NQA test group with the administrator name admin and the operation tag test.
[SwitchA] nqa entry admin test
# Configure the test type as ICMP-echo.
[SwitchA-nqa-admin-test] type icmp-echo
# Configure the destination address as 10.1.2.2.
[SwitchA-nqa-admin-test-icmp-echo] destination ip 10.1.2.2
# Set the test frequency to 100 ms.
[SwitchA-nqa-admin-test-icmp-echo] frequency 100
# Configure reaction entry 1, specifying that five consecutive probe failures trigger the track module.
[SwitchA-nqa-admin-test-icmp-echo] reaction 1 checked-element probe-fail
threshold-type consecutive 5 action-type trigger-only
[SwitchA-nqa-admin-test-icmp-echo] quit
# Start the NQA test.
[SwitchA] nqa schedule admin test start-time now lifetime forever
```

3. Configure a track entry on Switch A:

```
# Configure track entry 1, and associate it with reaction entry 1 of the NQA test group (with the administrator admin, and the operation tag test).
```

```
[SwitchA] track 1 nqa entry admin test reaction 1
```

4. Configure VRRP on Switch A:

```
# Create VRRP group 1, and configure the virtual IP address 10.1.1.10 for the group.
```

```
[SwitchA] interface vlan-interface 2
```

```
[SwitchA-Vlan-interface2] vrrp vrid 1 virtual-ip 10.1.1.10
```

```
# Set the priority of Switch A in VRRP group 1 to 110.
```

```
[SwitchA-Vlan-interface2] vrrp vrid 1 priority 110
```



```

# Set the authentication mode of VRRP group 1 to simple, and the authentication key to hello.
[SwitchA-Vlan-interface2] vrrp vrid 1 authentication-mode simple hello
# Configure the master to send VRRP packets at an interval of five seconds.
[SwitchA-Vlan-interface2] vrrp vrid 1 timer advertise 5
# Configure Switch A to operate in preemptive mode, and set the preemption delay to five
seconds.
[SwitchA-Vlan-interface2] vrrp vrid 1 preempt-mode timer delay 5
# Configure to monitor track entry 1, and specify the priority decrement to 30.
[SwitchA-Vlan-interface2] vrrp vrid 1 track 1 reduced 30

```

5. Configure VRRP on Switch B:

```

<SwitchB> system-view
[SwitchB] interface vlan-interface 2
# Create VRRP group 1, and configure the virtual IP address 10.1.1.10 for the group.
[SwitchB-Vlan-interface2] vrrp vrid 1 virtual-ip 10.1.1.10
# Set the authentication mode of VRRP group 1 to simple, and the authentication key to hello.
[SwitchB-Vlan-interface2] vrrp vrid 1 authentication-mode simple hello
# Configure the master to send VRRP packets at an interval of five seconds.
[SwitchB-Vlan-interface2] vrrp vrid 1 timer advertise 5
# Configure Switch B to operate in preemptive mode, and set the preemption delay to five
seconds.
[SwitchB-Vlan-interface2] vrrp vrid 1 preempt-mode timer delay 5

```

6. Verify the configuration:

After configuration, ping Host B on Host A, and you can see that Host B is reachable. Use the **display vrrp** command to view the configuration result.

Display detailed information about VRRP group 1 on Switch A.

```

[SwitchA-Vlan-interface2] display vrrp verbose
IPv4 Standby Information:
    Run Mode      : Standard
    Run Method    : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
    VRID          : 1                Adver Timer  : 5
    Admin Status  : Up                State         : Master
    Config Pri    : 110               Running Pri   : 110
    Preempt Mode  : Yes                Delay Time    : 5
    Auth Type     : Simple              Key           : hello
    Virtual IP    : 10.1.1.10
    Virtual MAC   : 0000-5e00-0101
    Master IP     : 10.1.1.1
VRRP Track Information:
    Track Object  : 1                State : Positive      Pri Reduced : 30

```

Display detailed information about VRRP group 1 on Switch B.

```

[SwitchB-Vlan-interface2] display vrrp verbose
IPv4 Standby Information:
    Run Mode      : Standard
    Run Method    : Virtual MAC

```

```

Total number of virtual routers : 1
Interface Vlan-interface2
  VRID          : 1                Adver Timer   : 5
  Admin Status  : Up                State          : Backup
  Config Pri    : 100               Running Pri    : 100
  Preempt Mode  : Yes               Delay Time     : 5
  Become Master : 2200ms left
  Auth Type     : Simple            Key            : hello
  Virtual IP    : 10.1.1.10
  Master IP     : 10.1.1.1

```

The output shows that in VRRP group 1, Switch A is the master, and Switch B is a backup. Packets from Host A to Host B are forwarded through Switch A.

When a fault is on the link between Switch A and Switch C, you can still successfully ping Host B on Host A. Use the **display vrrp** command to view information about VRRP group 1.

Display detailed information about VRRP group 1 on Switch A when a fault is on the link between Switch A and Switch C.

```

IPv4 Standby Information:
  Run Mode       : Standard
  Run Method     : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
  VRID          : 1                Adver Timer   : 5
  Admin Status  : Up                State          : Backup
  Config Pri    : 110               Running Pri    : 80
  Preempt Mode  : Yes               Delay Time     : 5
  Become Master : 2200ms left
  Auth Type     : Simple            Key            : hello
  Virtual IP    : 10.1.1.10
  Master IP     : 10.1.1.2
VRRP Track Information:
  Track Object   : 1                State : Negative          Pri Reduced : 30

```

Display detailed information about VRRP group 1 on Switch B when a fault is on the link between Switch A and Switch C.

```

[SwitchB-Vlan-interface2] display vrrp verbose
IPv4 Standby Information:
  Run Mode       : Standard
  Run Method     : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
  VRID          : 1                Adver Timer   : 5
  Admin Status  : Up                State          : Master
  Config Pri    : 100               Running Pri    : 100
  Preempt Mode  : Yes               Delay Time     : 5
  Auth Type     : Simple            Key            : hello
  Virtual IP    : 10.1.1.10
  Virtual MAC   : 0000-5e00-0101
  Master IP     : 10.1.1.2

```

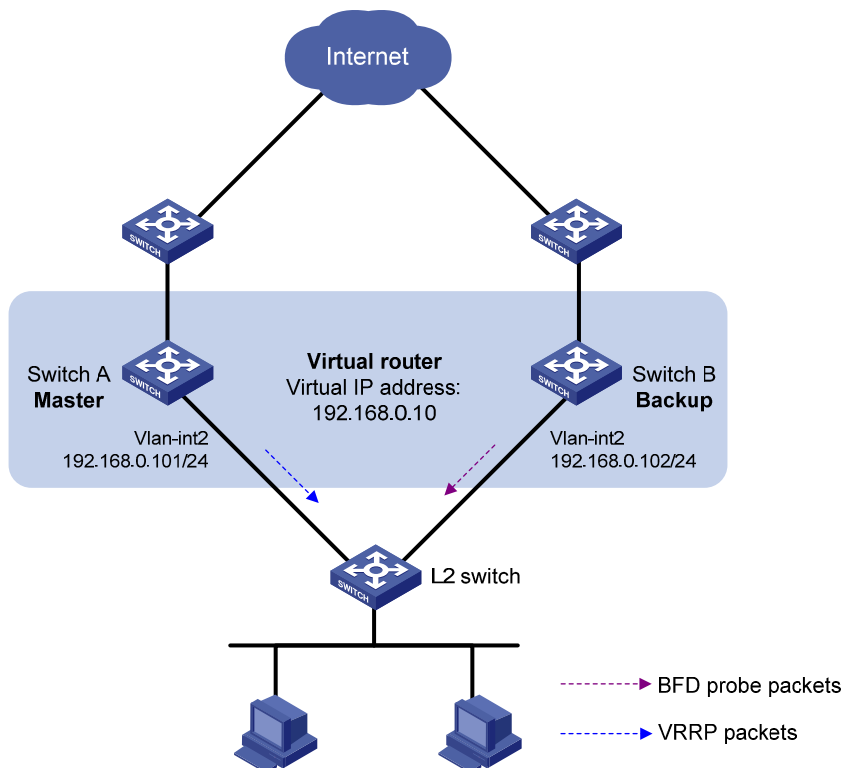
The output shows that when a fault is on the link between Switch A and Switch C, the priority of Switch A decreases to 80. Switch A becomes the backup, and Switch B becomes the master. Packets from Host A to Host B are forwarded through Switch B.

Configuring BFD for a VRRP backup to monitor the master (available only on the HP 5500 EI)

Network requirements

- As shown in Figure 57, Switch A and Switch B belong to VRRP group 1, whose virtual IP address is 192.168.0.10.
- The default gateway of the hosts in the LAN is 192.168.0.10. When Switch A works properly, the hosts in the LAN access the external network through Switch A. When Switch A fails, the hosts in the LAN access the external network through Switch B.
- If BFD is not configured, when the master in a VRRP group fails, the backup cannot become the master until the configured timeout timer expires. The timeout is generally three to four seconds, which makes the switchover slow. To solve this problem, VRRP uses BFD to probe the state of the master. Once the master fails, the backup can become the new master in milliseconds.

Figure 57 Network diagram



Configuration procedure

1. Create VLANs, and assign corresponding ports to the VLANs. Configure the IP address of each VLAN interface as shown in Figure 57. (Details not shown.)
2. Configure VRRP on Switch A:

```
<SwitchA> system-view  
[SwitchA] interface vlan-interface 2
```

Create VRRP group 1, and configure the virtual IP address 192.168.0.10 for the group. Set the priority of Switch A in VRRP group 1 to 110.

```
[SwitchA-Vlan-interface2] vrrp vrid 1 virtual-ip 192.168.0.10
[SwitchA-Vlan-interface2] vrrp vrid 1 priority 110
[SwitchA-Vlan-interface2] return
```

3. Configure BFD on Switch B:

Configure the source address of BFD echo packets as 10.10.10.10.

```
<SwitchB> system-view
[SwitchB] bfd echo-source-ip 10.10.10.10
```

4. Create the track entry to be associated with the BFD session on Switch B:

Create track entry 1 to be associated with the BFD session to check whether Switch A is reachable.

```
[SwitchB] track 1 bfd echo interface vlan-interface 2 remote ip 192.168.0.101 local
ip 192.168.0.102
```

5. Configure VRRP on Switch B:

Create VRRP group 1, and configure the virtual IP address 192.168.0.10 for the group. VRRP group 1 monitors the status of track entry 1. When the status of the track entry becomes Negative, Switch B quickly becomes the master.

```
[SwitchB] interface vlan-interface 2
[SwitchB-Vlan-interface2] vrrp vrid 1 virtual-ip 192.168.0.10
[SwitchB-Vlan-interface2] vrrp vrid 1 track 1 switchover
[SwitchB-Vlan-interface2] return
```

6. Verify the configuration:

Display the detailed information about VRRP group 1 on Switch A.

```
<SwitchA> display vrrp verbose
IPv4 Standby Information:
  Run Mode      : Standard
  Run Method    : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
  VRID          : 1                Adver Timer  : 1
  Admin Status  : Up              State         : Master
  Config Pri    : 110            Running Pri   : 110
  Preempt Mode  : Yes            Delay Time    : 0
  Auth Type     : None
  Virtual IP    : 192.168.0.10
  Virtual MAC   : 0000-5e00-0101
  Master IP     : 192.168.0.101
```

Display the detailed information about VRRP group 1 on Switch B.

```
<SwitchB> display vrrp verbose
IPv4 Standby Information:
  Run Mode      : Standard
  Run Method    : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
  VRID          : 1                Adver Timer  : 1
  Admin Status  : Up              State         : Backup
```

```
Config Pri      : 100           Running Pri    : 100
Preempt Mode    : Yes           Delay Time     : 0
Become Master   : 2200ms left
Auth Type       : None
Virtual IP      : 192.168.0.10
Master IP       : 192.168.0.101
```

VRRP Track Information:

```
Track Object    : 1             State : Positive           Switchover
```

Display information about track entry 1 on Switch B.

```
<SwitchB> display track 1
```

Track ID: 1

Status: Positive

Duration: 0 days 0 hours 0 minutes 32 seconds

Notification delay: Positive 0, Negative 0 (in seconds)

Reference object:

BFD session:

Packet type: Echo

Interface : Vlan-interface2

Remote IP : 192.168.0.101

Local IP : 192.168.0.102

The output shows that when the status of the track entry becomes Positive, Switch A is the master, and Switch B the backup.

Enable VRRP state debugging and BFD event debugging on Switch B.

```
<SwitchB> terminal debugging
```

```
<SwitchB> terminal monitor
```

```
<SwitchB> debugging vrrp state
```

```
<SwitchB> debugging bfd event
```

When Switch A fails, the following output is displayed on Switch B.

```
*Dec 17 14:44:34:142 2008 SwitchB BFD/7/EVENT:Send sess-down Msg,
[Src:192.168.0.102,Dst:192.168.0.101,Vlan-interface2,Echo], instance:0,
protocol:Track
```

```
*Dec 17 14:44:34:144 2008 SwitchB VRRP/7/DebugState: IPv4 Vlan-interface2 | Virtual
Router 1 : Backup --> Master reason: The status of the tracked object changed
```

Display the detailed information about the VRRP group on Switch B.

```
<SwitchB> display vrrp verbose
```

IPv4 Standby Information:

```
Run Mode       : Standard
```

```
Run Method     : Virtual MAC
```

Total number of virtual routers : 1

Interface Vlan-interface2

```
VRID          : 1             Adver Timer   : 1
```

```
Admin Status  : Up           State          : Master
```

```
Config Pri    : 100          Running Pri   : 100
```

```
Preempt Mode  : Yes          Delay Time    : 0
```

```
Auth Type     : None
```

```
Virtual IP    : 192.168.0.10
```

```
Virtual MAC   : 0000-5e00-0101
```

```
Master IP     : 192.168.0.102
```

VRRP Track Information:

Track Object : 1 State : Negative Switchover

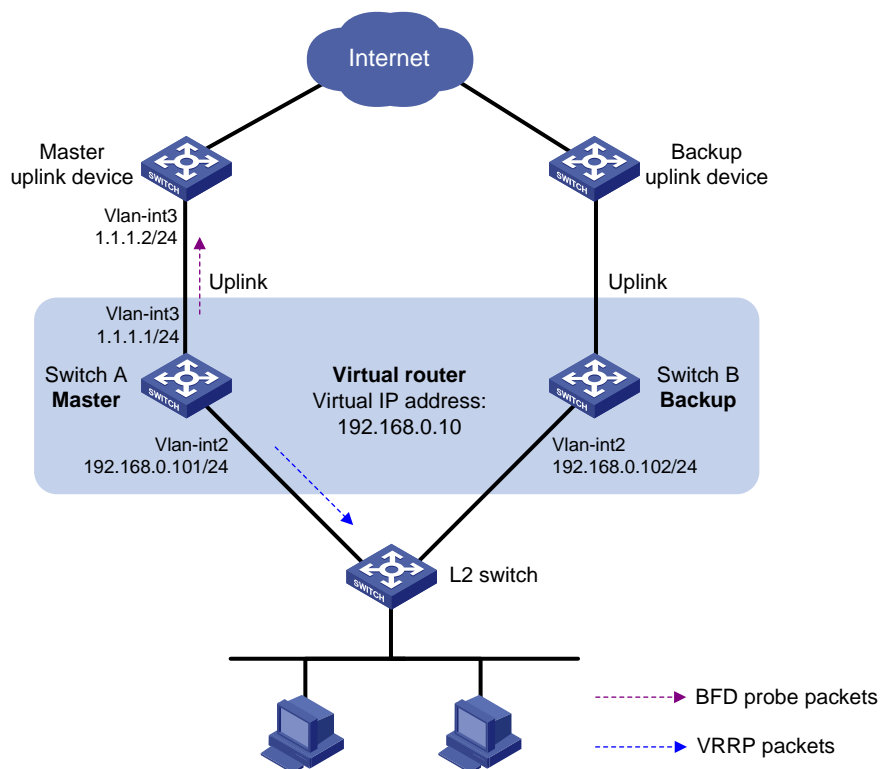
The output shows that when BFD detects that Switch A fails, it notifies VRRP through the track module to change the status of Switch B to master without waiting for a period three times the advertisement interval, so that a backup can quickly preempt as the master.

Configuring BFD for the VRRP master to monitor the uplinks (available only on the HP 5500 EI)

Network requirements

- As shown in Figure 58, Switch A and Switch B belong to VRRP group 1, whose virtual IP address is 192.168.0.10.
- The default gateway of the hosts in the LAN is 192.168.0.10.
- When Switch A works properly, the hosts in the LAN access the external network through Switch A. When Switch A detects that the uplink is down through BFD, it decreases its priority so that Switch B can preempt as the master, ensuring that the hosts in the LAN can access the external network through Switch B.

Figure 58 Network diagram



Configuration procedure

1. Create VLANs, and assign corresponding ports to the VLANs. Configure the IP address of each VLAN interface as shown in Figure 58. (Details not shown.)
2. Configure BFD on Switch A:
Configure the source address of BFD echo packets as 10.10.10.10.
<SwitchA> system-view

```
[SwitchA] bfd echo-source-ip 10.10.10.10
```

3. Create a track entry to be associated with the BFD session on Switch A:

```
# Create track entry 1 to be associated with the BFD session to check whether the uplink device with the IP address 1.1.1.2 is reachable.
```

```
[SwitchA] track 1 bfd echo interface vlan-interface 3 remote ip 1.1.1.2 local ip 1.1.1.1
```

4. Configure VRRP on Switch A:

```
# Create VRRP group 1, and configure the virtual IP address of the group as 192.168.0.10. Configure the priority of Switch A in VRRP group 1 as 110. Configure VRRP group 1 to monitor the status of track entry 1. When the status of the track entry becomes Negative, the priority of Switch A decreases by 20.
```

```
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] vrrp vrid 1 virtual-ip 192.168.0.10
[SwitchA-Vlan-interface2] vrrp vrid 1 priority 110
[SwitchA-Vlan-interface2] vrrp vrid 1 track 1 reduced 20
[SwitchA-Vlan-interface2] return
```

5. Configure VRRP on Switch B:

```
# Create VRRP group 1, and configure the virtual IP address of the group as 192.168.0.10.
```

```
<SwitchB> system-view
[SwitchB] interface vlan-interface 2
[SwitchB-Vlan-interface2] vrrp vrid 1 virtual-ip 192.168.0.10
[SwitchB-Vlan-interface2] return
```

6. Verify the configuration:

```
# Display the detailed information about the VRRP group on Switch A.
```

```
<SwitchA> display vrrp verbose
IPv4 Standby Information:
  Run Mode      : Standard
  Run Method    : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
  VRID          : 1                Adver Timer   : 1
  Admin Status  : Up              State          : Master
  Config Pri    : 110             Running Pri    : 110
  Preempt Mode  : Yes             Delay Time     : 0
  Auth Type     : None
  Virtual IP    : 192.168.0.10
  Virtual MAC   : 0000-5e00-0101
  Master IP     : 192.168.0.101
VRRP Track Information:
  Track Object  : 1                State : Positive  Pri Reduced : 20
```

```
# Display the information about track entry 1 on Switch A.
```

```
<SwitchA> display track 1
Track ID: 1
  Status: Positive
  Duration: 0 days 0 hours 0 minutes 32 seconds
  Notification delay: Positive 0, Negative 0 (in seconds)
  Reference object:
```

```
BFD session:
Packet type: Echo
Interface   : Vlan-interface2
Remote IP   : 1.1.1.2
Local IP    : 1.1.1.1
```

Display the detailed information about the VRRP group on Switch B.

```
<SwitchB> display vrrp verbose
IPv4 Standby Information:
  Run Mode      : Standard
  Run Method    : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
  VRID          : 1                Adver Timer  : 1
  Admin Status  : Up              State         : Backup
  Config Pri    : 100            Running Pri   : 100
  Preempt Mode  : Yes            Delay Time    : 0
  Become Master : 2200ms left
  Auth Type     : None
  Virtual IP    : 192.168.0.10
  Master IP     : 192.168.0.101
```

The output shows that when the status of track entry 1 becomes **Positive**, Switch A is the master, and Switch B the backup.

When the uplink of Switch A goes down, the status of track entry 1 becomes **Negative**.

```
<SwitchA> display track 1
Track ID: 1
  Status: Negative
  Duration: 0 days 0 hours 0 minutes 32 seconds
  Notification delay: Positive 0, Negative 0 (in seconds)
  Reference object:
    BFD session:
      Packet type: Echo
      Interface   : Vlan-interface2
      Remote IP   : 1.1.1.2
      Local IP    : 1.1.1.1
```

Display the detailed information about VRRP group 1 on Switch A.

```
<SwitchA> display vrrp verbose
IPv4 Standby Information:
  Run Mode      : Standard
  Run Method    : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
  VRID          : 1                Adver Timer  : 1
  Admin Status  : Up              State         : Backup
  Config Pri    : 110            Running Pri   : 90
  Preempt Mode  : Yes            Delay Time    : 0
  Become Master : 2200ms left
  Auth Type     : None
  Virtual IP    : 192.168.0.10
```



```

Master IP      : 192.168.0.102
VRRP Track Information:
Track Object   : 1                State : Negative   Pri Reduced : 20
# Display the detailed information about VRRP group 1 on Switch B.
<SwitchB> display vrrp verbose
IPv4 Standby Information:
Run Mode      : Standard
Run Method    : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
VRID          : 1                Adver Timer   : 1
Admin Status  : Up              State          : Master
Config Pri    : 100             Running Pri    : 100
Preempt Mode  : Yes             Delay Time     : 0
Auth Type     : None
Virtual IP    : 192.168.0.10
Virtual MAC   : 0000-5e00-0101
Master IP     : 192.168.0.102

```

The output shows that when Switch A detects that the uplink fails through BFD, it decreases its priority by 20 to make sure that Switch B can preempt as the master.

Static routing-track-NQA collaboration configuration example

Network requirements

As shown in [Figure 59](#), Switch A, Switch B, Switch C, and Switch D are connected to two segments 20.1.1.0/24 and 30.1.1.0/24. Configure static routes on these switches so that the two segments can communicate with each other, and configure route backup to improve reliability of the network.

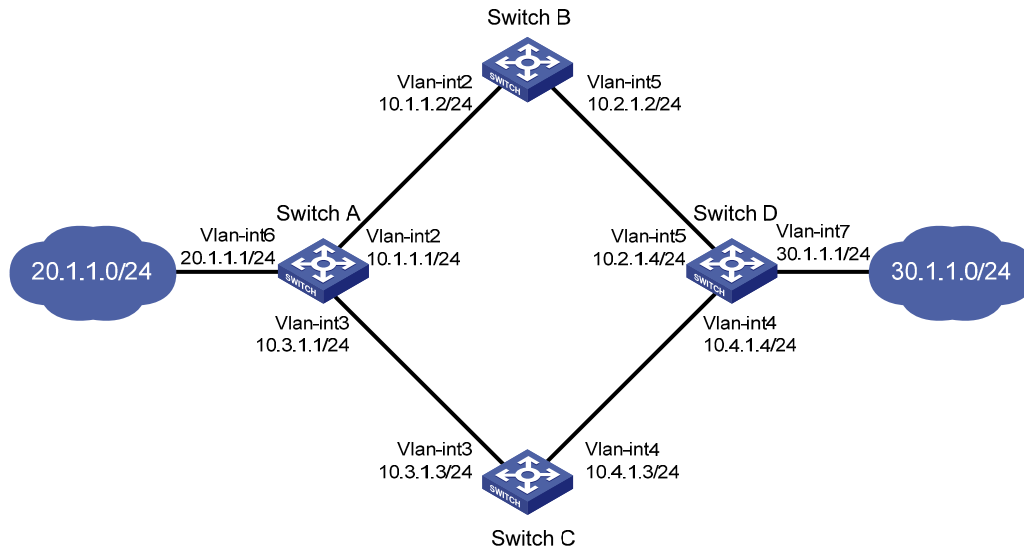
Switch A is the default gateway of the hosts in segment 20.1.1.0/24. Two static routes to 30.1.1.0/24 exist on Switch A, with the next hop being Switch B and Switch C, respectively. These two static routes back up each other as follows:

- The static route with Switch B as the next hop has a higher priority, and is the master route. If this route is available, Switch A forwards packets to 30.1.1.0/24 through Switch B.
- The static route with Switch C as the next hop acts as the backup route.
- Configure static routing-track-NQA collaboration to determine whether the master route is available in real time. If the master route is unavailable, the backup route takes effect, and Switch A forwards packets to 30.1.1.0/24 through Switch C.

Similarly, Switch D is the default gateway of the hosts in segment 30.1.1.0/24. Two static routes to 20.1.1.0/24 exist on Switch D, with the next hop being Switch B and Switch C, respectively. These two static routes back up each other as follows:

- The static route with Switch B as the next hop has a higher priority, and is the master route. If this route is available, Switch D forwards packets to 20.1.1.0/24 through Switch B.
- The static route with Switch C as the next hop acts as the backup route.
- Configure static routing-track-NQA collaboration to determine whether the master route is available in real time. If the master route is unavailable, the backup route takes effect, and Switch D forwards packets to 20.1.1.0/24 through Switch C.

Figure 59 Network diagram



Configuration procedure

1. Create VLANs, and assign corresponding ports to the VLANs. Configure the IP address of each VLAN interface as shown in Figure 59. (Details not shown.)
2. Configure Switch A:

Configure a static route to 30.1.1.0/24, with the address of the next hop as 10.1.1.2 and the default priority 60. This static route is associated with track entry 1.

```
<SwitchA> system-view
```

```
[SwitchA] ip route-static 30.1.1.0 24 10.1.1.2 track 1
```

Configure a static route to 30.1.1.0/24, with the address of the next hop as 10.3.1.3 and the priority 80.

```
[SwitchA] ip route-static 30.1.1.0 24 10.3.1.3 preference 80
```

Configure a static route to 10.2.1.4, with the address of the next hop as 10.1.1.2.

```
[SwitchA] ip route-static 10.2.1.4 24 10.1.1.2
```

Create an NQA test group with the administrator **admin** and the operation tag **test**.

```
[SwitchA] nqa entry admin test
```

Configure the test type as ICMP-echo.

```
[SwitchA-nqa-admin-test] type icmp-echo
```

Configure the destination address of the test as 10.2.1.4 and the next hop address as 10.1.1.2 to check the connectivity of the path from Switch A to Switch B and then to Switch D through NQA.

```
[SwitchA-nqa-admin-test-icmp-echo] destination ip 10.2.1.4
```

```
[SwitchA-nqa-admin-test-icmp-echo] next-hop 10.1.1.2
```

Configure the test frequency as 100 ms.

```
[SwitchA-nqa-admin-test-icmp-echo] frequency 100
```

Configure reaction entry 1, specifying that five consecutive probe failures trigger the track module.

```
[SwitchA-nqa-admin-test-icmp-echo] reaction 1 checked-element probe-fail
```

```
threshold-type consecutive 5 action-type trigger-only
```

```
[SwitchA-nqa-admin-test-icmp-echo] quit
```

Start the NQA test.

```
[SwitchA] nqa schedule admin test start-time now lifetime forever
```

```
# Configure track entry 1, and associate it with reaction entry 1 of the NQA test group (with the administrator admin, and the operation tag test).
```

```
[SwitchA] track 1 nqa entry admin test reaction 1
```

3. Configure Switch B:

```
# Configure a static route to 30.1.1.0/24, with the address of the next hop as 10.2.1.4.
```

```
<SwitchB> system-view
```

```
[SwitchB] ip route-static 30.1.1.0 24 10.2.1.4
```

```
# Configure a static route to 20.1.1.0/24, with the address of the next hop as 10.1.1.1.
```

```
[SwitchB] ip route-static 20.1.1.0 24 10.1.1.1
```

4. Configure Switch C:

```
# Configure a static route to 30.1.1.0/24, with the address of the next hop as 10.4.1.4.
```

```
<SwitchC> system-view
```

```
[SwitchC] ip route-static 30.1.1.0 24 10.4.1.4
```

```
# Configure a static route to 20.1.1.0/24, with the address of the next hop as 10.3.1.1.
```

```
[SwitchC] ip route-static 20.1.1.0 24 10.3.1.1
```

5. Configure Switch D:

```
# Configure a static route to 20.1.1.0/24, with the address of the next hop as 10.2.1.2 and the default priority 60. This static route is associated with track entry 1.
```

```
<SwitchD> system-view
```

```
[SwitchD] ip route-static 20.1.1.0 24 10.2.1.2 track 1
```

```
# Configure a static route to 20.1.1.0/24, with the address of the next hop as 10.4.1.3 and the priority 80.
```

```
[SwitchD] ip route-static 20.1.1.0 24 10.4.1.3 preference 80
```

```
# Configure a static route to 10.1.1.1, with the address of the next hop as 10.2.1.2.
```

```
[SwitchD] ip route-static 10.1.1.1 24 10.2.1.2
```

```
# Create an NQA test group with the administrator admin and the operation tag test.
```

```
[SwitchD] nqa entry admin test
```

```
# Configure the test type as ICMP-echo.
```

```
[SwitchD-nqa-admin-test] type icmp-echo
```

```
# Configure the destination address of the test as 10.1.1.1 and the next hop address as 10.2.1.2 to check the connectivity of the path from Switch D to Switch B and then to Switch A through NQA.
```

```
[SwitchD-nqa-admin-test-icmp-echo] destination ip 10.1.1.1
```

```
[SwitchD-nqa-admin-test-icmp-echo] next-hop 10.2.1.2
```

```
# Configure the test frequency as 100 ms.
```

```
[SwitchD-nqa-admin-test-icmp-echo] frequency 100
```

```
# Configure reaction entry 1, specifying that five consecutive probe failures trigger the track module.
```

```
[SwitchD-nqa-admin-test-icmp-echo] reaction 1 checked-element probe-fail
```

```
threshold-type consecutive 5 action-type trigger-only
```

```
[SwitchD-nqa-admin-test-icmp-echo] quit
```

```
# Start the NQA test.
```

```
[SwitchD] nqa schedule admin test start-time now lifetime forever
```

```
# Configure track entry 1, and associate it with reaction entry 1 of the NQA test group (with the administrator admin, and the operation tag test).
```

```
[SwitchD] track 1 nqa entry admin test reaction 1
```

6. Verify the configuration:

Display information about the track entry on Switch A.

```
[SwitchA] display track all
```

```
Track ID: 1
  Status: Positive
  Duration: 0 days 0 hours 0 minutes 32 seconds
  Notification delay: Positive 0, Negative 0 (in seconds)
  Reference object:
    NQA entry: admin test
    Reaction: 1
```

Display the routing table of Switch A.

```
[SwitchA] display ip routing-table
```

```
Routing Tables: Public
```

```
          Destinations : 10          Routes : 10
Destination/Mask    Proto Pre  Cost           NextHop         Interface
10.1.1.0/24         Direct 0    0             10.1.1.1        Vlan2
10.1.1.1/32         Direct 0    0             127.0.0.1       InLoop0
10.2.1.0/24         Static 60   0             10.1.1.2        Vlan2
10.3.1.0/24         Direct 0    0             10.3.1.1        Vlan3
10.3.1.1/32         Direct 0    0             127.0.0.1       InLoop0
20.1.1.0/24         Direct 0    0             20.1.1.1        Vlan6
20.1.1.1/32         Direct 0    0             127.0.0.1       InLoop0
30.1.1.0/24         Static 60   0             10.1.1.2        Vlan2
127.0.0.0/8         Direct 0    0             127.0.0.1       InLoop0
127.0.0.1/32        Direct 0    0             127.0.0.1       InLoop0
```

The output shows the NQA test result: the master route is available (the status of the track entry is Positive), and Switch A forwards packets to 30.1.1.0/24 through Switch B.

Remove the IP address of interface VLAN-interface 2 on Switch B.

```
<SwitchB> system-view
```

```
[SwitchB] interface vlan-interface 2
```

```
[SwitchB-Vlan-interface2] undo ip address
```

Display information about the track entry on Switch A.

```
[SwitchA] display track all
```

```
Track ID: 1
  Status: Negative
  Duration: 0 days 0 hours 0 minutes 32 seconds
  Notification delay: Positive 0, Negative 0 (in seconds)
  Reference object:
    NQA entry: admin test
    Reaction: 1
```

Display the routing table of Switch A.

```
[SwitchA] display ip routing-table
```

```
Routing Tables: Public
```

```
          Destinations : 10          Routes : 10
```

| Destination/Mask | Proto | Pre | Cost | NextHop | Interface |
|------------------|--------|-----|------|-----------|-----------|
| 10.1.1.0/24 | Direct | 0 | 0 | 10.1.1.1 | Vlan2 |
| 10.1.1.1/32 | Direct | 0 | 0 | 127.0.0.1 | InLoop0 |
| 10.2.1.0/24 | Static | 60 | 0 | 10.1.1.2 | Vlan2 |
| 10.3.1.0/24 | Direct | 0 | 0 | 10.3.1.1 | Vlan3 |
| 10.3.1.1/32 | Direct | 0 | 0 | 127.0.0.1 | InLoop0 |
| 20.1.1.0/24 | Direct | 0 | 0 | 20.1.1.1 | Vlan6 |
| 20.1.1.1/32 | Direct | 0 | 0 | 127.0.0.1 | InLoop0 |
| 30.1.1.0/24 | Static | 80 | 0 | 10.3.1.3 | Vlan3 |
| 127.0.0.0/8 | Direct | 0 | 0 | 127.0.0.1 | InLoop0 |
| 127.0.0.1/32 | Direct | 0 | 0 | 127.0.0.1 | InLoop0 |

The output shows the NQA test result: the master route is unavailable (the status of the track entry is Negative). The backup static route takes effect and Switch A forwards packets to 30.1.1.0/24 through Switch C.

When the master route fails, the hosts in 20.1.1.0/24 can still communicate with the hosts in 30.1.1.0/24.

```
[SwitchA] ping -a 20.1.1.1 30.1.1.1
PING 30.1.1.1: 56 data bytes, press CTRL_C to break
  Reply from 30.1.1.1: bytes=56 Sequence=1 ttl=254 time=2 ms
  Reply from 30.1.1.1: bytes=56 Sequence=2 ttl=254 time=1 ms
  Reply from 30.1.1.1: bytes=56 Sequence=3 ttl=254 time=1 ms
  Reply from 30.1.1.1: bytes=56 Sequence=4 ttl=254 time=2 ms
  Reply from 30.1.1.1: bytes=56 Sequence=5 ttl=254 time=1 ms
--- 30.1.1.1 ping statistics ---
  5 packet(s) transmitted
  5 packet(s) received
  0.00% packet loss
  round-trip min/avg/max = 1/1/2 ms
```

The output on Switch D is similar to that on Switch A. When the master route fails, the hosts in 30.1.1.0/24 can still communicate with the hosts in 20.1.1.0/24.

```
[SwitchB] ping -a 30.1.1.1 20.1.1.1
PING 20.1.1.1: 56 data bytes, press CTRL_C to break
  Reply from 20.1.1.1: bytes=56 Sequence=1 ttl=254 time=2 ms
  Reply from 20.1.1.1: bytes=56 Sequence=2 ttl=254 time=1 ms
  Reply from 20.1.1.1: bytes=56 Sequence=3 ttl=254 time=1 ms
  Reply from 20.1.1.1: bytes=56 Sequence=4 ttl=254 time=1 ms
  Reply from 20.1.1.1: bytes=56 Sequence=5 ttl=254 time=1 ms
--- 20.1.1.1 ping statistics ---
  5 packet(s) transmitted
  5 packet(s) received
  0.00% packet loss
  round-trip min/avg/max = 1/1/2 ms
```

Static routing-Track-BFD collaboration configuration example (available only on the HP 5500 EI)

Network requirements

As shown in [Figure 60](#), Switch A, Switch B, and Switch C are connected to two segments 20.1.1.0/24 and 30.1.1.0/24. Configure static routes on these routers so that the two segments can communicate with each other, and configure route backup to improve reliability of the network.

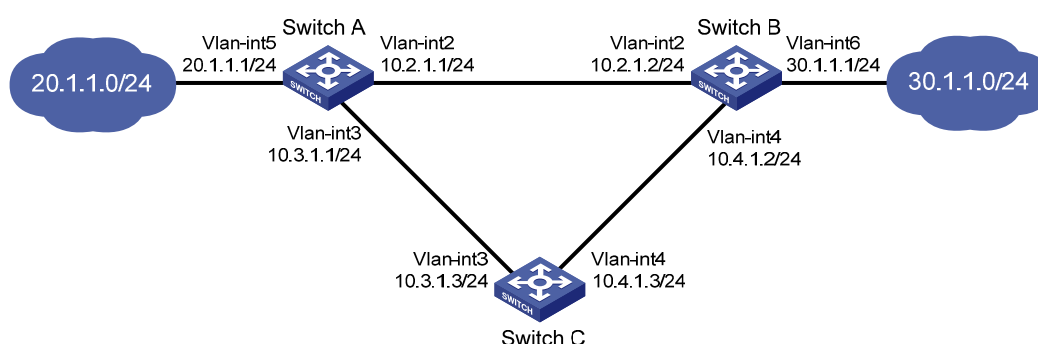
Switch A is the default gateway of the hosts in segment 20.1.1.0/24. Two static routes to 30.1.1.0/24 exist on Switch A, with the next hop being Switch B and Switch C, respectively. These two static routes back up each other as follows:

- The static route with Switch B as the next hop has a higher priority and is the master route. If this route is available, Switch A forwards packets to 30.1.1.0/24 through Switch B.
- The static route with Switch C as the next hop acts as the backup route.
- Configure static routing-track-BFD collaboration to determine whether the master route is available in real time. If the master route is unavailable, BFD can quickly detect the route failure to make the backup route take effect, and Switch A forwards packets to 30.1.1.0/24 through Switch C and Switch B.

Similarly, Switch B is the default gateway of the hosts in segment 30.1.1.0/24. Two static routes to 20.1.1.0/24 exist on Switch B, with the next hop being Switch A and Switch C, respectively. These two static routes back up each other as follows:

- The static route with Switch A as the next hop has a higher priority and is the master route. If this route is available, Switch B forwards packets to 20.1.1.0/24 through Switch A.
- The static route with Switch C as the next hop acts as the backup route.
- Configure static routing-track-BFD collaboration to determine whether the master route is available in real time. If the master route is unavailable, BFD can quickly detect the route failure to make the backup route take effect, and Switch B forwards packets to 20.1.1.0/24 through Switch C and Switch A.

Figure 60 Network diagram



Configuration procedure

1. Create VLANs, and assign corresponding ports to the VLANs. Configure the IP address of each VLAN interface as shown in [Figure 60](#). (Details not shown.)
2. Configure Switch A:
Configure a static route to 30.1.1.0/24, with the address of the next hop as 10.2.1.2 and the default priority 60. This static route is associated with track entry 1.

```

<SwitchA> system-view
[SwitchA] ip route-static 30.1.1.0 24 10.2.1.2 track 1
# Configure a static route to 30.1.1.0/24, with the address of the next hop as 10.3.1.3 and the
priority 80.
[SwitchA] ip route-static 30.1.1.0 24 10.3.1.3 preference 80
# Configure the source address of BFD echo packets as 10.10.10.10.
[SwitchA] bfd echo-source-ip 10.10.10.10
# Configure track entry 1, and associate it with the BFD session. Check whether Switch A can be
interoperated with the next hop of static route (Switch B).
[SwitchA] track 1 bfd echo interface vlan-interface 2 remote ip 10.2.1.2 local ip
10.2.1.1

```

3. Configure Switch B:

```

# Configure a static route to 20.1.1.0/24, with the address of the next hop as 10.2.1.1 and the
default priority 60. This static route is associated with track entry 1.
<SwitchB> system-view
[SwitchB] ip route-static 20.1.1.0 24 10.2.1.1 track 1
# Configure a static route to 20.1.1.0/24, with the address of the next hop as 10.4.1.3 and the
priority 80.
[SwitchB] ip route-static 20.1.1.0 24 10.4.1.3 preference 80
# Configure the source address of BFD echo packets as 1.1.1.1.
[SwitchB] bfd echo-source-ip 1.1.1.1
# Configure track entry 1 that is associated with the BFD session to check whether Switch B can
communicate with the next hop (Switch A) of the static route.
[SwitchB] track 1 bfd echo interface vlan-interface 2 remote ip 10.2.1.1 local ip
10.2.1.2

```

4. Configure Switch C:

```

# Configure a static route to 30.1.1.0/24, with the address of the next hop as 10.4.1.2.
<SwitchC> system-view
[SwitchC] ip route-static 30.1.1.0 24 10.4.1.2
# Configure a static route to 20.1.1.0/24, with the address of the next hop as 10.3.1.1.
[SwitchB] ip route-static 20.1.1.0 24 10.3.1.1

```

5. Verify the configuration:

```

# Display information about the track entry on Switch A.
[SwitchA] display track all
Track ID: 1
  Status: Positive
  Duration: 0 days 0 hours 0 minutes 32 seconds
  Notification delay: Positive 0, Negative 0 (in seconds)
  Reference object:
    BFD Session:
      Packet type: Echo
      Interface   : Vlan-interface2
      Remote IP   : 10.2.1.2
      Local IP    : 10.2.1.1
# Display the routing table of Switch A.
[SwitchA] display ip routing-table

```

Routing Tables: Public

| Destination/Mask | Proto | Pre | Cost | NextHop | Interface |
|------------------|--------|-----|------|-----------|-----------|
| 10.2.1.0/24 | Direct | 0 | 0 | 10.2.1.1 | Vlan2 |
| 10.2.1.1/32 | Direct | 0 | 0 | 127.0.0.1 | InLoop0 |
| 10.3.1.0/24 | Direct | 0 | 0 | 10.3.1.1 | Vlan3 |
| 10.3.1.1/32 | Direct | 0 | 0 | 127.0.0.1 | InLoop0 |
| 20.1.1.0/24 | Direct | 0 | 0 | 20.1.1.1 | Vlan5 |
| 20.1.1.1/32 | Direct | 0 | 0 | 127.0.0.1 | InLoop0 |
| 30.1.1.0/24 | Static | 60 | 0 | 10.2.1.2 | Vlan2 |
| 127.0.0.0/8 | Direct | 0 | 0 | 127.0.0.1 | InLoop0 |
| 127.0.0.1/32 | Direct | 0 | 0 | 127.0.0.1 | InLoop0 |

The output shows the BFD detection result: the next hop 10.2.1.2 is reachable (the status of the track entry is Positive). The master static route takes effect. Switch A forwards packets to 30.1.1.0/24 through Switch B.

Remove the IP address of interface VLAN-interface 2 on Switch B.

```
<SwitchB> system-view
[SwitchB] interface vlan-interface 2
[SwitchB-Vlan-interface2] undo ip address
```

Display information about the track entry on Switch A.

```
[SwitchA] display track all
Track ID: 1
  Status: Negative
  Duration: 0 days 0 hours 0 minutes 32 seconds
  Notification delay: Positive 0, Negative 0 (in seconds)
  Reference object:
    BFD Session:
      Packet type: Echo
      Interface   : Vlan-interface2
      Remote IP   : 10.2.1.2
      Local IP    : 10.2.1.1
```

Display the routing table of Switch A.

```
[SwitchA] display ip routing-table
Routing Tables: Public
          Destinations : 9          Routes : 9

Destination/Mask    Proto Pre  Cost      NextHop         Interface
10.2.1.0/24         Direct 0    0         10.2.1.1        Vlan2
10.2.1.1/32         Direct 0    0         127.0.0.1       InLoop0
10.3.1.0/24         Direct 0    0         10.3.1.1        Vlan3
10.3.1.1/32         Direct 0    0         127.0.0.1       InLoop0
20.1.1.0/24         Direct 0    0         20.1.1.1        Vlan5
20.1.1.1/32         Direct 0    0         127.0.0.1       InLoop0
30.1.1.0/24         Static 80   0         10.3.1.3        Vlan3
127.0.0.0/8         Direct 0    0         127.0.0.1       InLoop0
127.0.0.1/32        Direct 0    0         127.0.0.1       InLoop0
```

The output shows the BFD detection result: the next hop 10.2.1.2 is unreachable (the status of the track entry is Negative), and the backup static route takes effect, and Switch A forwards packets to 30.1.1.0/24 through Switch C and Switch B.

When the master route fails, the hosts in 20.1.1.0/24 can still communicate with the hosts in 30.1.1.0/24.

```
[SwitchA] ping -a 20.1.1.1 30.1.1.1
PING 30.1.1.1: 56 data bytes, press CTRL_C to break
  Reply from 30.1.1.1: bytes=56 Sequence=1 ttl=254 time=2 ms
  Reply from 30.1.1.1: bytes=56 Sequence=2 ttl=254 time=1 ms
  Reply from 30.1.1.1: bytes=56 Sequence=3 ttl=254 time=1 ms
  Reply from 30.1.1.1: bytes=56 Sequence=4 ttl=254 time=2 ms
  Reply from 30.1.1.1: bytes=56 Sequence=5 ttl=254 time=1 ms
--- 30.1.1.1 ping statistics ---
  5 packet(s) transmitted
  5 packet(s) received
  0.00% packet loss
  round-trip min/avg/max = 1/1/2 ms
```

The output on Switch B is similar to that on Switch A. When the master route fails, the hosts in 30.1.1.0/24 can still communicate with the hosts in 20.1.1.0/24.

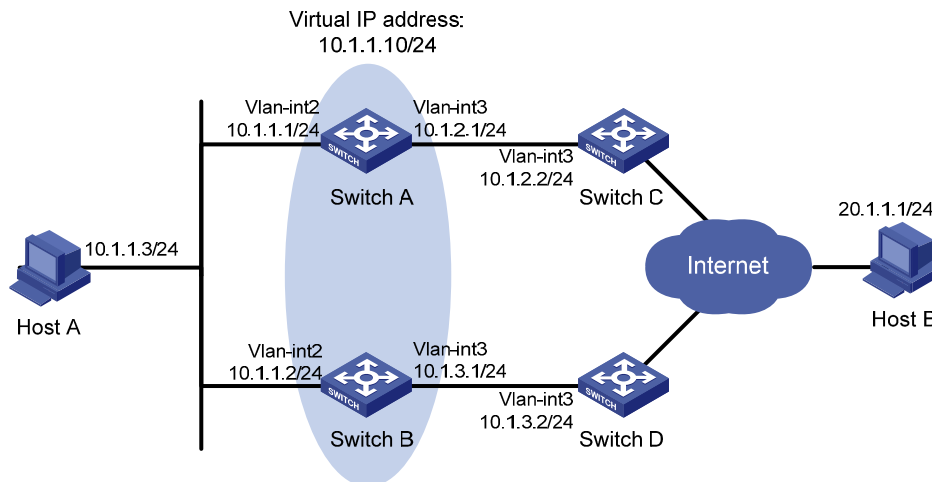
```
[SwitchB] ping -a 30.1.1.1 20.1.1.1
PING 20.1.1.1: 56 data bytes, press CTRL_C to break
  Reply from 20.1.1.1: bytes=56 Sequence=1 ttl=254 time=2 ms
  Reply from 20.1.1.1: bytes=56 Sequence=2 ttl=254 time=1 ms
  Reply from 20.1.1.1: bytes=56 Sequence=3 ttl=254 time=1 ms
  Reply from 20.1.1.1: bytes=56 Sequence=4 ttl=254 time=1 ms
  Reply from 20.1.1.1: bytes=56 Sequence=5 ttl=254 time=1 ms
--- 20.1.1.1 ping statistics ---
  5 packet(s) transmitted
  5 packet(s) received
  0.00% packet loss
  round-trip min/avg/max = 1/1/2 ms
```

VRRP-track-interface management collaboration configuration example (the master monitors the uplink interface) (available only on the HP 5500 EI)

Network requirements

- As shown in [Figure 61](#), Host A needs to access Host B on the Internet. The default gateway of Host A is 10.1.1.10/24.
- Switch A and Switch B belong to VRRP group 1, whose virtual IP address is 10.1.1.10.
- When Switch A works properly, packets from Host A to Host B are forwarded through Switch A. When VRRP detects that a fault is on the uplink interface of Switch A through the interface management module, packets from Host A to Host B are forwarded through Switch B.

Figure 61 Network diagram



Configuration procedure

1. Create VLANs, and assign corresponding ports to the VLANs. Configure the IP address of each VLAN interface as shown in Figure 61. (Details not shown.)

2. Configure a track entry on Switch A:

```
# Configure track entry 1, and associate it with the physical status of the uplink interface  
Vlan-interface 3.
```

```
[SwitchA] track 1 interface vlan-interface 3
```

3. Configure VRRP on Switch A:

```
# Create VRRP group 1, and configure the virtual IP address 10.1.1.10 for the group.
```

```
[SwitchA] interface vlan-interface 2
```

```
[SwitchA-Vlan-interface2] vrrp vrid 1 virtual-ip 10.1.1.10
```

```
# Set the priority of Switch A in VRRP group 1 to 110.
```

```
[SwitchA-Vlan-interface2] vrrp vrid 1 priority 110
```

```
# Configure to monitor track entry 1, and specify the priority decrement as 30.
```

```
[SwitchA-Vlan-interface2] vrrp vrid 1 track 1 reduced 30
```

4. Configure VRRP on Switch B:

```
<SwitchB> system-view
```

```
[SwitchB] interface vlan-interface 2
```

```
# Create VRRP group 1, and configure the virtual IP address 10.1.1.10 for the group.
```

```
[SwitchB-Vlan-interface2] vrrp vrid 1 virtual-ip 10.1.1.10
```

5. Verify the configuration:

After configuration, ping Host B on Host A, and you can see that Host B is reachable. Use the **display vrrp** command to view the configuration result.

```
# Display detailed information about VRRP group 1 on Switch A.
```

```
[SwitchA-Vlan-interface2] display vrrp verbose
```

```
IPv4 Standby Information:
```

```
Run Mode      : Standard
```

```
Run Method    : Virtual MAC
```

```
Total number of virtual routers : 1
```

```
Interface Vlan-interface2
```

```

VRID          : 1                Adver Timer   : 1
Admin Status  : Up                State        : Master
Config Pri    : 110               Running Pri   : 110
Preempt Mode  : Yes               Delay Time    : 0
Auth Type     : None
Virtual IP    : 10.1.1.10
Virtual MAC   : 0000-5e00-0101
Master IP     : 10.1.1.1
VRRP Track Information:
Track Object  : 1                State : Positive  Pri Reduced : 30

```

Display detailed information about VRRP group 1 on Switch B.

```

[SwitchB-Vlan-interface2] display vrrp verbose
IPv4 Standby Information:
Run Mode      : Standard
Run Method    : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
VRID          : 1                Adver Timer   : 1
Admin Status  : Up                State        : Backup
Config Pri    : 100               Running Pri   : 100
Preempt Mode  : Yes               Delay Time    : 0
Become Master : 2200ms left
Auth Type     : None
Virtual IP    : 10.1.1.10
Master IP     : 10.1.1.1

```

The output shows that in VRRP group 1, Switch A is the master, and Switch B is a backup. Packets from Host A to Host B are forwarded through Switch A.

Shut down the uplink interface VLAN-interface 3 on Switch A.

```

[SwitchA-Vlan-interface2] interface vlan-interface 3
[SwitchA-Vlan-interface3] shutdown

```

After shutting down the uplink interface on Switch A, you can still successfully ping Host B on Host A. Use the **display vrrp** command to view information about VRRP group 1.

After shutting down the uplink interface on Switch A, display detailed information about VRRP group 1 on Switch A.

```

[SwitchA-Vlan-interface3] display vrrp verbose
IPv4 Standby Information:
Run Mode      : Standard
Run Method    : Virtual MAC
Total number of virtual routers : 1
Interface Vlan-interface2
VRID          : 1                Adver Timer   : 1
Admin Status  : Up                State        : Backup
Config Pri    : 110               Running Pri   : 80
Preempt Mode  : Yes               Delay Time    : 0
Become Master : 2200ms left
Auth Type     : None
Virtual IP    : 10.1.1.10
Master IP     : 10.1.1.2

```

VRRP Track Information:

Track Object : 1 State : Negative Pri Reduced : 30

After shutting down the uplink interface on Switch A, display detailed information about VRRP group 1 on Switch B.

```
[SwitchB-Vlan-interface2] display vrrp verbose
```

IPv4 Standby Information:

Run Mode : Standard

Run Method : Virtual MAC

Total number of virtual routers : 1

Interface Vlan-interface2

VRID : 1 Adver Timer : 1

Admin Status : Up State : Master

Config Pri : 100 Running Pri : 100

Preempt Mode : Yes Delay Time : 0

Auth Type : None

Virtual IP : 10.1.1.10

Virtual MAC : 0000-5e00-0101

Master IP : 10.1.1.2

The output shows that when the uplink interface on Switch A is shut down, the priority of Switch A decreases to 80. Switch A becomes the backup, and Switch B becomes the master. Packets from Host A to Host B are forwarded through Switch B.

Index

[A](#) [B](#) [C](#) [D](#) [E](#) [H](#) [I](#) [M](#) [R](#) [S](#) [T](#) [V](#)

A

- Activating an RRPP domain, [67](#)
- Associating the track module with a detection module, [202](#)
- Associating the track module with an application module, [204](#)
- Availability evaluation, [1](#)
- Availability requirements, [1](#)

B

- BFD overview, [193](#)

C

- CFD configuration example, [29](#)
- CFD configuration task list, [20](#)
- CFD overview, [16](#)
- Configuration guidelines, [192](#)
- Configuring a Smart Link device, [100](#)
- Configuring an associated device, [104](#)
- Configuring an RRPP ring group, [68](#)
- Configuring basic CFD settings, [21](#)
- Configuring basic Ethernet OAM functions, [9](#)
- Configuring BFD basic functions, [198](#)
- Configuring CFD functions, [24](#)
- Configuring control VLANs, [63](#)
- Configuring DLDAP authentication, [45](#)
- Configuring Ethernet OAM remote loopback, [11](#)
- Configuring link monitoring, [10](#)
- Configuring Monitor Link, [120](#)
- Configuring protected VLANs, [64](#)
- Configuring RRPP rings, [65](#)
- Configuring RRPP timers, [68](#)
- Configuring the backup VLAN, [190](#)
- Configuring the duplex mode and speed of an Ethernet interface, [42](#)
- Configuring the Ethernet OAM connection detection timers, [9](#)
- Configuring VRRP for IPv4, [136](#)
- Configuring VRRP for IPv6, [143](#)

- Creating an RRPP domain, [63](#)

D

- Displaying and maintaining BFD, [199](#)
- Displaying and maintaining CFD, [28](#)
- Displaying and maintaining DLDAP, [46](#)
- Displaying and maintaining Ethernet OAM configuration, [13](#)
- Displaying and maintaining Monitor Link, [121](#)
- Displaying and maintaining RRPP, [69](#)
- Displaying and maintaining Smart Link, [104](#)
- Displaying and maintaining stateful failover, [191](#)
- Displaying and maintaining track entries, [208](#)
- DLDAP configuration examples, [46](#)
- DLDAP configuration task list, [42](#)
- DLDAP overview, [35](#)

E

- Enabling DLDAP, [43](#)
- Enabling stateful failover, [190](#)
- Ethernet OAM configuration example, [13](#)
- Ethernet OAM configuration task list, [8](#)
- Ethernet OAM overview, [5](#)

H

- High availability technologies, [2](#)

I

- Introduction to stateful failover configuration, [189](#)
- IPv4-based VRRP configuration examples, [148](#)
- IPv6-based VRRP configuration examples, [167](#)

M

- Monitor Link configuration example, [121](#)
- Monitor Link overview, [119](#)

R

- Resetting DLDAP state, [45](#)
- RRPP configuration examples, [69](#)
- RRPP configuration task list, [62](#)
- RRPP overview, [54](#)

S

- Setting DLDP mode, [43](#)
- Setting the delaydown timer, [44](#)
- Setting the interval to send advertisement packets, [43](#)
- Setting the port shutdown mode, [44](#)
- Smart Link configuration examples, [105](#)
- Smart Link configuration task list, [100](#)
- Smart Link overview, [97](#)
- Stateful failover configuration example, [191](#)
- Stateful failover overview, [188](#)

T

- Track configuration examples, [208](#)
- Track configuration task list, [201](#)
- Track overview, [200](#)
- Troubleshooting, [96](#)
- Troubleshooting DLDP, [53](#)
- Troubleshooting VRRP, [187](#)

V

- VRRP load balancing mode, [131](#)
- VRRP overview, [125](#)
- VRRP standard protocol mode, [126](#)

Contents

| | |
|--|----|
| Using ping, tracer, and system debugging | 1 |
| Ping | 1 |
| Using a ping command to test network connectivity | 1 |
| Ping example | 1 |
| Tracer | 3 |
| Prerequisites | 4 |
| Using a tracer command to identify failed or all nodes in a path | 5 |
| System debugging | 5 |
| Debugging information control switches | 5 |
| Debugging a feature module | 6 |
| Ping and tracer example | 7 |
| Configuring NTP | 8 |
| Overview | 8 |
| NTP application | 8 |
| NTP advantages | 8 |
| How NTP works | 9 |
| NTP message format | 10 |
| Operation modes | 11 |
| NTP configuration task list | 13 |
| Configuring NTP operation modes | 13 |
| Configuring the client/server mode | 14 |
| Configuring the symmetric peers mode | 14 |
| Configuring the broadcast mode | 15 |
| Configuring the multicast mode | 16 |
| Configuring optional parameters | 17 |
| Specifying the source interface for NTP messages | 17 |
| Disabling an interface from receiving NTP messages | 17 |
| Configuring the allowed maximum number of dynamic sessions | 18 |
| Configuring the DSCP value for NTP messages | 18 |
| Configuring access-control rights | 18 |
| Configuration prerequisites | 19 |
| Configuration procedure | 19 |
| Configuring NTP authentication | 19 |
| Configuring NTP authentication in client/server mode | 19 |
| Displaying and maintaining NTP | 21 |
| NTP configuration examples | 21 |
| Configuring the client/server mode | 21 |
| Configuring the NTP symmetric mode | 22 |
| Configuring NTP broadcast mode | 24 |
| Configuring NTP multicast mode | 25 |
| Configuring NTP client/server mode with authentication | 28 |
| Configuring NTP broadcast mode with authentication | 29 |
| Configuring the information center | 33 |
| Overview | 33 |
| Classification of system information | 34 |
| System information levels | 34 |
| System information channels and output destinations | 34 |
| Outputting system information by source module | 35 |

| | |
|---|-----------|
| Default output rules of system information | 35 |
| System information format | 36 |
| Information center configuration task list | 39 |
| Outputting system information to the console | 40 |
| Configuring a system information output rule for the console | 40 |
| Enabling system information output to the console | 40 |
| Outputting system information to the monitor terminal | 41 |
| Configuring a system information output rule for the monitor terminal | 41 |
| Enabling system information output to the monitor terminal | 42 |
| Outputting system information to a log host | 42 |
| Outputting system information to the trap buffer | 43 |
| Outputting system information to the log buffer | 43 |
| Outputting system information to the SNMP module | 44 |
| Outputting system information to the Web interface | 45 |
| Saving security logs into the security log file | 45 |
| Saving security logs into the security log file | 46 |
| Managing the security log file | 47 |
| Configuring synchronous information output | 49 |
| Disabling an interface from generating link up/down logging information | 49 |
| Displaying and maintaining information center | 50 |
| Information center configuration examples | 50 |
| Outputting log information to a UNIX log host | 50 |
| Outputting log information to a Linux log host | 51 |
| Outputting log information to the console | 53 |
| Saving security logs into the security log file | 54 |
| Configuring SNMP | 57 |
| Overview | 57 |
| SNMP framework | 57 |
| MIB and view-based MIB access control | 57 |
| SNMP operations | 58 |
| SNMP protocol versions | 58 |
| SNMP configuration task list | 58 |
| Configuring SNMP basic parameters | 58 |
| Configuring SNMPv3 basic parameters | 58 |
| Configuring SNMPv1 or SNMPv2c basic parameters | 60 |
| Switching the NM-specific interface index format | 62 |
| Configuration guidelines | 62 |
| Configuration procedure | 62 |
| Configuring SNMP logging | 63 |
| Configuring SNMP traps | 63 |
| Enabling SNMP traps | 63 |
| Configuring the SNMP agent to send traps to a host | 64 |
| Displaying and maintaining SNMP | 65 |
| SNMP configuration examples | 66 |
| SNMPv1/SNMPv2c configuration example | 66 |
| SNMPv3 configuration example | 67 |
| SNMP logging configuration example | 69 |
| Configuring RMON | 71 |
| Overview | 71 |
| Working mechanism | 71 |
| RMON groups | 71 |
| Configuring the RMON statistics function | 73 |
| Configuring the RMON Ethernet statistics function | 73 |

| | |
|---|------------|
| Configuring the RMON history statistics function | 73 |
| Configuring the RMON alarm function | 74 |
| Displaying and maintaining RMON | 75 |
| Ethernet statistics group configuration example | 76 |
| History group configuration example | 76 |
| Alarm group configuration example | 78 |
| Configuring port mirroring | 81 |
| Introduction to port mirroring | 81 |
| Terminologies of port mirroring | 81 |
| Port mirroring classification and implementation | 82 |
| Configuring local port mirroring | 84 |
| Local port mirroring configuration task list | 84 |
| Creating a local mirroring group | 84 |
| Configuring source ports for the local mirroring group | 84 |
| Configuring the monitor port for the local mirroring group | 85 |
| Using the remote probe VLAN to enable local mirroring to support multiple monitor ports | 86 |
| Configuring layer 2 remote port mirroring | 87 |
| Layer 2 remote port mirroring configuration task list | 87 |
| Configuring a remote source group (on the source device) | 88 |
| Configuring a remote destination group (on the destination device) | 90 |
| Displaying and maintaining port mirroring | 92 |
| Port mirroring configuration examples | 92 |
| Local port mirroring configuration example | 92 |
| Local port mirroring with multiple monitor ports configuration example | 93 |
| Layer 2 remote port mirroring configuration example | 94 |
| Configuring traffic mirroring | 97 |
| Introduction to traffic mirroring | 97 |
| Traffic mirroring configuration task list | 97 |
| Configuring match criteria | 97 |
| Configuring traffic mirroring of different types | 98 |
| Mirroring traffic to a port | 98 |
| Mirroring traffic to the CPU | 98 |
| Configuring a QoS policy | 98 |
| Applying a QoS policy | 99 |
| Apply a QoS policy to a port | 99 |
| Apply a QoS policy to a VLAN | 99 |
| Apply a QoS policy globally | 100 |
| Apply a QoS policy to the control plane | 100 |
| Displaying and maintaining traffic mirroring | 100 |
| Traffic mirroring configuration example | 101 |
| Traffic mirroring configuration example | 101 |
| Configuring NQA | 103 |
| Overview | 103 |
| NQA features | 103 |
| NQA concepts | 105 |
| NQA probe operation procedure | 106 |
| NQA configuration task list | 106 |
| Configuring the NQA server | 107 |
| Enabling the NQA client | 107 |
| Creating an NQA test group | 108 |
| Configuring an NQA test group | 108 |
| Configuring ICMP echo tests | 108 |
| Configuring DHCP tests | 109 |

| | |
|---|------------|
| Configuring DNS tests | 110 |
| Configuring FTP tests | 111 |
| Configuring HTTP tests | 112 |
| Configuring UDP jitter tests | 112 |
| Configuring SNMP tests | 114 |
| Configuring TCP tests | 115 |
| Configuring UDP echo tests | 116 |
| Configuring voice tests | 116 |
| Configuring DLSw tests | 118 |
| Configuring the collaboration function | 119 |
| Configuring threshold monitoring | 120 |
| Configuration prerequisites | 120 |
| Configuration guidelines | 120 |
| Configuration procedure | 120 |
| Configuring the NQA statistics collection function | 122 |
| Configuring the history records saving function | 122 |
| Configuring optional parameters for an NQA test group | 123 |
| Configuring a schedule for an NQA test group | 124 |
| Configuration prerequisites | 124 |
| Configuration guidelines | 125 |
| Configuration procedure | 125 |
| Displaying and maintaining NQA | 125 |
| NQA configuration examples | 126 |
| ICMP echo test configuration example | 126 |
| DHCP test configuration example | 127 |
| DNS test configuration example | 129 |
| FTP test configuration example | 130 |
| HTTP test configuration example | 131 |
| UDP jitter test configuration example | 132 |
| SNMP test configuration example | 135 |
| TCP test configuration example | 136 |
| UDP echo test configuration example | 137 |
| Voice test configuration example | 139 |
| DLSw test configuration example | 141 |
| NQA collaboration configuration example | 143 |
| Configuring sFlow | 146 |
| sFlow overview | 146 |
| Introduction to sFlow | 146 |
| sFlow operation | 147 |
| Configuring sFlow | 147 |
| Configuring the sFlow agent and sFlow collector | 147 |
| Configuring flow sampling | 148 |
| Configuring counter sampling | 148 |
| Displaying and maintaining sFlow | 148 |
| sFlow configuration example | 149 |
| Network requirements | 149 |
| Troubleshooting sFlow configuration | 150 |
| Configuring IPC | 151 |
| Overview | 151 |
| Node | 151 |
| Link | 151 |
| Channel | 151 |
| Packet sending modes | 152 |

| | |
|--|------------|
| Enabling IPC performance statistics | 152 |
| Displaying and maintaining IPC | 153 |
| Configuring PoE | 154 |
| Overview | 154 |
| Protocol specification | 154 |
| PoE configuration task list | 155 |
| Configuration guidelines | 155 |
| Enabling PoE | 156 |
| Enabling PoE for a PoE interface | 156 |
| Detecting PDs | 157 |
| Enabling the PSE to detect nonstandard PDs | 157 |
| Configuring a PD disconnection detection mode | 157 |
| Configuring the PoE power | 157 |
| Configuring the maximum PoE interface power | 157 |
| Configuring PoE power management | 158 |
| Configuring PoE interface power management | 158 |
| Configuring the PoE monitoring function | 159 |
| Configuring PSE power monitoring | 159 |
| Monitoring PD | 159 |
| Configuring PoE interface through PoE profile | 159 |
| Configuration guidelines | 160 |
| Configuring PoE profile | 160 |
| Applying PoE profile | 160 |
| Upgrading PSE processing software in service | 161 |
| Configuration guidelines | 161 |
| Displaying and maintaining PoE | 161 |
| PoE configuration example | 162 |
| Troubleshooting PoE | 163 |
| Setting the priority of a PoE interface to critical fails | 163 |
| Applying a PoE profile to a PoE interface fails | 163 |
| Configuring an AC input under-voltage threshold fails | 164 |
| Configuring cluster management | 165 |
| Overview | 165 |
| Roles in a cluster | 165 |
| How a cluster works | 166 |
| Cluster management configuration task list | 169 |
| Configuration guidelines | 169 |
| Configuring the management switch | 170 |
| Enabling NDP globally and for specific ports | 170 |
| Configuring NDP parameters | 171 |
| Enabling NTDP globally and for specific ports | 171 |
| Configuring NTDP parameters | 172 |
| Manually collecting topology information | 172 |
| Enabling the cluster function | 173 |
| Establishing a cluster | 173 |
| Enabling management VLAN auto-negotiation | 174 |
| Configuring communication between the management switch and the member switches within a cluster | 174 |
| Configuring cluster management protocol packets | 175 |
| Cluster member management | 176 |
| Configuring the member switches | 177 |
| Enabling NDP | 177 |
| Enabling NTDP | 177 |

| | |
|--|------------|
| Manually collecting topology information | 177 |
| Enabling the cluster function | 177 |
| Deleting a member switch from a cluster | 177 |
| Configuring access between the management switch and its member switches | 177 |
| Configuration guidelines | 177 |
| Configuration procedure | 178 |
| Adding a candidate switch to a cluster | 178 |
| Configuring advanced cluster management functions | 178 |
| Configuring topology management | 178 |
| Configuring interaction for a cluster | 179 |
| SNMP configuration synchronization function | 180 |
| Configuring web user accounts in batches | 181 |
| Displaying and maintaining cluster management | 181 |
| Cluster management configuration example | 182 |
| Network requirements | 182 |
| Configuration procedure | 183 |
| Configuring CWMP | 186 |
| Hardware compatibility | 186 |
| CWMP overview | 186 |
| CWMP network framework | 186 |
| Basic CWMP functions | 187 |
| CWMP mechanism | 188 |
| CWMP configuration tasks | 191 |
| Configuring the DHCP server | 191 |
| Configuring the DNS server | 192 |
| Configuring the ACS server | 192 |
| Configuring CPEs | 192 |
| Enabling CWMP | 192 |
| Configuring the ACS attributes | 193 |
| Configuring the ACS URL | 193 |
| Configuring the ACS username and password | 193 |
| Configuring CPE attributes | 194 |
| Configuring the CPE username and password | 194 |
| Configuring the CWMP connection interface | 194 |
| Sending Inform messages | 195 |
| Configuring the maximum number of attempts made to retry a connection | 195 |
| Configuring the close-wait timer of the CPE | 195 |
| Configuring the CPE working mode | 196 |
| Specifying an SSL client policy for HTTPS connection to ACS | 196 |
| Displaying and maintaining CWMP | 197 |
| CWMP configuration example | 197 |
| Configuration guidelines | 197 |
| Network requirements | 197 |
| Configuration procedure | 198 |
| Configuring a stack (available only on the HP 5500 SI) | 206 |
| Stack configuration task list | 206 |
| Configuring the master device of a stack | 206 |
| Configuring a private IP address pool for the stack | 207 |
| Configuring stack ports | 207 |
| Creating a stack | 207 |
| Configuring the stack ports of a member device | 207 |
| Logging in to the CLI of a member from the master | 208 |
| Displaying and maintaining stack configuration | 208 |

| | |
|-----------------------------------|-----|
| Stack configuration example | 208 |
| Index | 211 |

Using ping, tracert, and system debugging

Use the ping, tracert, and system debugging utilities to test network connectivity and identify network problems.

Ping

The ping utility sends ICMP echo requests (ECHO-REQUEST) to the destination device. Upon receiving the requests, the destination device responds with ICMP echo replies (ECHO-REPLY) to the source device. The source device outputs statistics about the ping operation, including the number of packets sent, number of echo replies received, and the round-trip time. You can measure the network performance by analyzing these statistics.

Using a ping command to test network connectivity

| Task | Command | Remarks |
|--|--|---|
| Check whether a specified address in an IP network is reachable. | <ul style="list-style-type: none">For an IPv4 network:
ping [ip] [-a source-ip -c count -f -h ttl -i interface-type interface-number -m interval -n -p pad -q -r -s packet-size -t timeout -tos tos -v -vpn-instance vpn-instance-name] * hostFor an IPv6 network:
ping ipv6 [-a source-ipv6 -c count -m interval -s packet-size -t timeout -tos tos -vpn-instance vpn-instance-name] * host [-i interface-type interface-number] | Use one of the commands.
Available in any view.
Only the HP 5500 EI Switch Series supports the vpn-instance keyword. |

⚠ IMPORTANT:

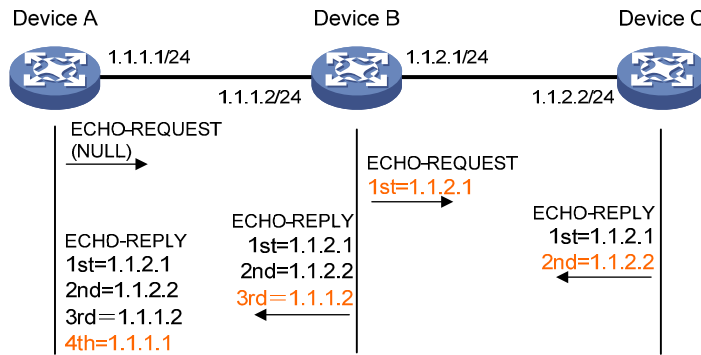
When you configure the **ping** command for a low-speed network, HP recommends that you set a larger value for the timeout timer (indicated by the **-t** keyword in the command).

Ping example

Network requirements

Test the network connectivity between Device A and Device C in [Figure 1](#). If they can reach each other, get detailed information about routes from Device A to Device C.

Figure 1 Network diagram



Test procedure

Use the **ping** command on Device A to test connectivity to Device C.

```
<DeviceA> ping 1.1.2.2
PING 1.1.2.2: 56 data bytes, press CTRL_C to break
  Reply from 1.1.2.2: bytes=56 Sequence=1 ttl=254 time=205 ms
  Reply from 1.1.2.2: bytes=56 Sequence=2 ttl=254 time=1 ms
  Reply from 1.1.2.2: bytes=56 Sequence=3 ttl=254 time=1 ms
  Reply from 1.1.2.2: bytes=56 Sequence=4 ttl=254 time=1 ms
  Reply from 1.1.2.2: bytes=56 Sequence=5 ttl=254 time=1 ms

--- 1.1.2.2 ping statistics ---
  5 packet(s) transmitted
  5 packet(s) received
  0.00% packet loss
round-trip min/avg/max = 1/41/205 ms
```

Get detailed information about routes from Device A to Device C.

```
<DeviceA> ping -r 1.1.2.2
PING 1.1.2.2: 56 data bytes, press CTRL_C to break
  Reply from 1.1.2.2: bytes=56 Sequence=1 ttl=254 time=53 ms
    Record Route:
      1.1.2.1
      1.1.2.2
      1.1.1.2
      1.1.1.1
  Reply from 1.1.2.2: bytes=56 Sequence=2 ttl=254 time=1 ms
    Record Route:
      1.1.2.1
      1.1.2.2
      1.1.1.2
      1.1.1.1
  Reply from 1.1.2.2: bytes=56 Sequence=3 ttl=254 time=1 ms
    Record Route:
      1.1.2.1
      1.1.2.2
      1.1.1.2
```

```

    1.1.1.1
Reply from 1.1.2.2: bytes=56 Sequence=4 ttl=254 time=1 ms
  Record Route:
    1.1.2.1
    1.1.2.2
    1.1.1.2
    1.1.1.1
Reply from 1.1.2.2: bytes=56 Sequence=5 ttl=254 time=1 ms
  Record Route:
    1.1.2.1
    1.1.2.2
    1.1.1.2
    1.1.1.1
--- 1.1.2.2 ping statistics ---
    5 packet(s) transmitted
    5 packet(s) received
    0.00% packet loss
round-trip min/avg/max = 1/11/53 ms

```

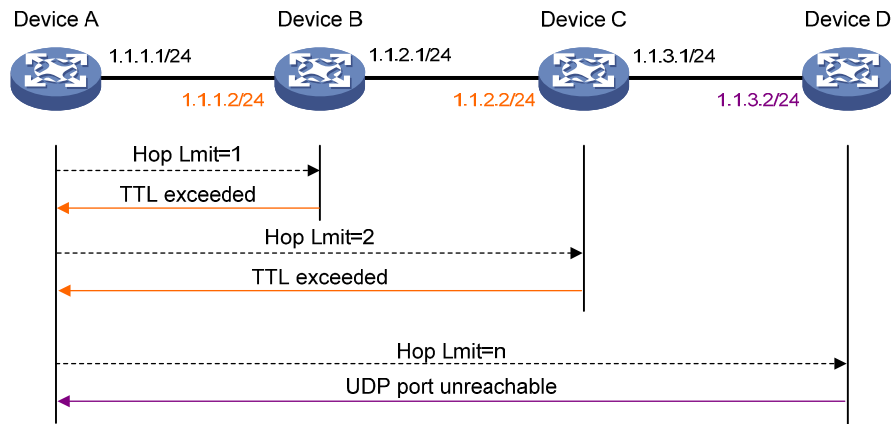
The test procedure with the **ping -r** command (see [Figure 1](#)) is as follows:

1. The source (Device A) sends an ICMP echo request with the RR option being empty to the destination (Device C).
2. The intermediate device (Device B) adds the IP address of its outbound interface (1.1.2.1) to the RR option of the ICMP echo request, and forwards the packet.
3. Upon receiving the request, the destination device copies the RR option in the request and adds the IP address of its outbound interface (1.1.2.2) to the RR option. Then the destination device sends an ICMP echo reply.
4. The intermediate device adds the IP address of its outbound interface (1.1.1.2) to the RR option in the ICMP echo reply, and then forwards the reply.
5. Upon receiving the reply, the source device adds the IP address of its inbound interface (1.1.1.1) to the RR option. Finally, you can get detailed information about routes from Device A to Device C: 1.1.1.1 <-> {1.1.1.2; 1.1.2.1} <-> 1.1.2.2.

Tracert

Tracert (also called "Traceroute") enables you to get the IP addresses of Layer 3 devices in the path to a specific destination. You can use tracert to test network connectivity and identify failed nodes.

Figure 2 Network diagram



Tracert uses received ICMP error messages to get the IP addresses of devices. As shown in [Figure 2](#), tracert works as follows:

1. The source device (Device A) sends a UDP packet with a TTL value of 1 to the destination device (Device D). The destination UDP port is not used by any application on the destination device.
2. The first hop (Device B, the first Layer 3 device that receives the packet) responds by sending a TTL-expired ICMP error message to the source, with its IP address encapsulated. In this way, the source device can get the address of the first Layer 3 device (1.1.1.2).
3. The source device sends a packet with a TTL value of 2 to the destination device.
4. The second hop (Device C) responds with a TTL-expired ICMP error message, which gives the source device the address of the second Layer 3 device (1.1.2.2).
5. The process continues until the packet sent by the source device reaches the ultimate destination device. Because no application uses the destination port specified in the packet, so the destination device responds with a port-unreachable ICMP message to the source device, with its IP address 1.1.3.2 encapsulated. This way, the source device gets the IP address of the destination device (1.1.3.2).
6. The source device thinks that the packet has reached the destination device after receiving the port-unreachable ICMP message, and the path to the destination device is 1.1.1.2 to 1.1.2.2 to 1.1.3.2.

Prerequisites

Before you use a tracert command, perform the tasks in this section.

For an IPv4 network:

- Enable sending of ICMP timeout packets on the intermediate devices (the devices between the source and destination devices). If the intermediate devices are HP devices, execute the **ip ttl-expires enable** command on the devices. For more information about this command, see *Layer 3—IP Services Command Reference*.
- Enable sending of ICMP destination unreachable packets on the destination device. If the destination device is an HP device, execute the **ip unreachable enable** command. For more information about this command, see *Layer 3—IP Services Command Reference*.

For an IPv6 network:

- Enable sending of ICMPv6 timeout packets on the intermediate devices (the devices between the source and destination devices). If the intermediate devices are HP devices, execute the **ipv6**

hoplimit-expires enable command on the devices. For more information about this command, see *Layer 3—IP Services Command Reference*.

- Enable sending of ICMPv6 destination unreachable packets on the destination device. If the destination device is an HP device, execute the **ipv6 unreachable enable** command. For more information about this command, see *Layer 3—IP Services Command Reference*.

Using a tracer command to identify failed or all nodes in a path

| Task | Command | Remarks |
|--|--|--|
| Display the routes from source to destination. | <ul style="list-style-type: none"> • For an IPv4 network:
 <code>tracer [-a source-ip -f first-ttl -m max-ttl -p port -q packet-number -tos tos -vpn-instance vpn-instance-name -w timeout] * host</code> • For an IPv6 network:
 <code>tracer ipv6 [-f first-ttl -m max-ttl -p port -q packet-number -tos tos -vpn-instance vpn-instance-name -w timeout] * host</code> | <p>Use one of the commands.</p> <p>Available in any view.</p> <p>Only the HP 5500 EI Switch Series supports the vpn-instance keyword.</p> |

System debugging

The device supports various debugging for the majority of protocols and features and provides debugging information to help users diagnose errors.

Debugging information control switches

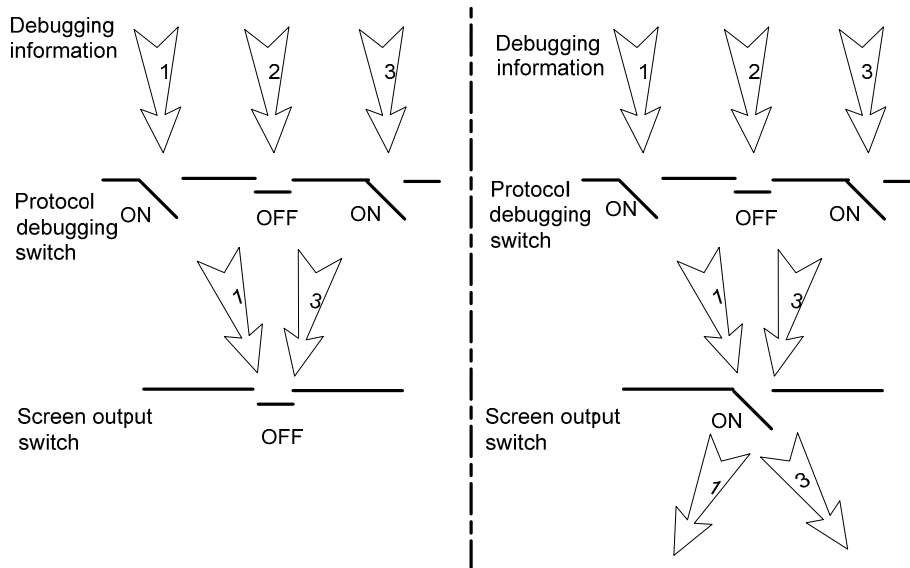
The following two switches control the display of debugging information:

- **Protocol debugging switch**—Controls protocol-specific debugging information.
- **Screen output switch**—Controls whether to display the debugging information on a certain screen.

As shown in [Figure 3](#), assume that the device can provide debugging for the three modules 1, 2, and 3. The debugging information can be output on a terminal only when both the protocol debugging switch and the screen output switch are turned on.

Output of debugging information depends on the configurations of the information center and the debugging commands of each protocol and functional module. Debugging information is typically displayed on a terminal (including console or VTY) for display. You can also send debugging information to other destinations. For more information, see "[Configuring the information center](#)."

Figure 3 Relationship between the protocol and screen output switch



Debugging a feature module

Output of debugging commands is memory intensive. To guarantee system performance, enable debugging only for modules that are in an exceptional condition. When debugging is complete, use the **undo debugging all** command to disable all the debugging functions.

Configure the **debugging**, **terminal debugging**, and **terminal monitor** commands before you can display detailed debugging information on the terminal. For more information about the **terminal debugging** and **terminal monitor** commands, see *Network Management and Monitoring Command Reference*.

To debug a feature module and display the debugging information on a terminal:

| Step | Command | Remarks |
|--|--|---|
| 1. Enable the terminal monitoring of system information. | terminal monitor | Optional.
The terminal monitoring on the console is enabled by default and the terminal monitoring on the monitoring terminal is disabled by default.
Available in user view. |
| 2. Enable the terminal display of debugging information. | terminal debugging | Disabled by default.
Available in user view. |
| 3. Enable debugging for a specified module. | debugging <i>module-name</i> [<i>option</i>] | Disabled by default.
Available in user view. |
| 4. Display the enabled debugging functions. | display debugging [interface <i>interface-type interface-number</i>] [<i>module-name</i>] [{ begin exclude include } <i>regular-expression</i>] | Optional.
Available in any view. |

Ping and traceroute example

Network requirements

As shown in Figure 4, Device A failed to Telnet Device C. Determine whether Device A and Device C can reach each other. If they cannot reach each other, locate the failed nodes in the network.

Figure 4 Network diagram



Test procedure

1. Use the **ping** command to test connectivity between Device A and Device C.

```
<DeviceA> ping 1.1.2.2
PING 1.1.2.2: 56 data bytes, press CTRL_C to break
Request time out
Request time out
Request time out
Request time out
Request time out

--- 1.1.2.2 ping statistics ---
 5 packet(s) transmitted
 0 packet(s) received
100.00% packet loss
```

The output shows that Device A and Device C cannot reach each other.

2. Use the **tracert** command to identify failed nodes.

```
<DeviceA> system-view
[DeviceA] ip ttl-expires enable
[DeviceA] ip unreachable enable
[DeviceA] tracert 1.1.2.2
  traceroute to 1.1.2.2(1.1.2.2) 30 hops max, 40 bytes packet, press CTRL_C to break
 1  1.1.1.2  14 ms  10 ms  20 ms
 2  * * *
 3  * * *
 4  * * *
 5
<DeviceA>
```

The output shows that Device A and Device C cannot reach each other, Device A and Device B can reach each other, and an error has occurred on the connection between Device B and Device C.

3. Use the **debugging ip icmp** command on Device A and Device C to verify that they can send and receive the specific ICMP packets, or use the **display ip routing-table** command to verify the availability of active routes between Device A and Device C.

Configuring NTP

The term *Layer 3 Ethernet port* refers to route-mode (or Layer 3) Ethernet ports. You can set an Ethernet port to operate in route mode by using the **port link-mode route** command (see *Layer 2—LAN Switching Configuration Guide*).

Overview

NTP is typically used in large networks to dynamically synchronize time among network devices. It guarantees higher clock accuracy than manual system clock setting. In a small network that does not require high clock accuracy, you can keep time synchronized among devices by changing their system clocks one by one.

NTP runs over UDP and uses UDP port 123.

NTP application

An administrator is unable to keep time synchronized among all devices within a network by changing the system clock on each station, because this is a huge work and does not guarantee clock precision. NTP, however, allows quick clock synchronization within the entire network and ensures high clock precision.

NTP is used when all devices within the network must keep consistent time. For example:

- In analyzing log and debugging information collected from different devices in network management, time must be used as a reference basis.
- All devices must use the same reference clock in a charging system.
- To implement certain functions, such as a scheduled restart of all devices within the network, all devices must keep consistent time.
- If multiple systems process a complex event in cooperation, these systems must use the same reference clock to ensure the correct execution sequence.
- For incremental backup between a backup server and clients, timekeeping must be synchronized between the backup server and all clients.

NTP advantages

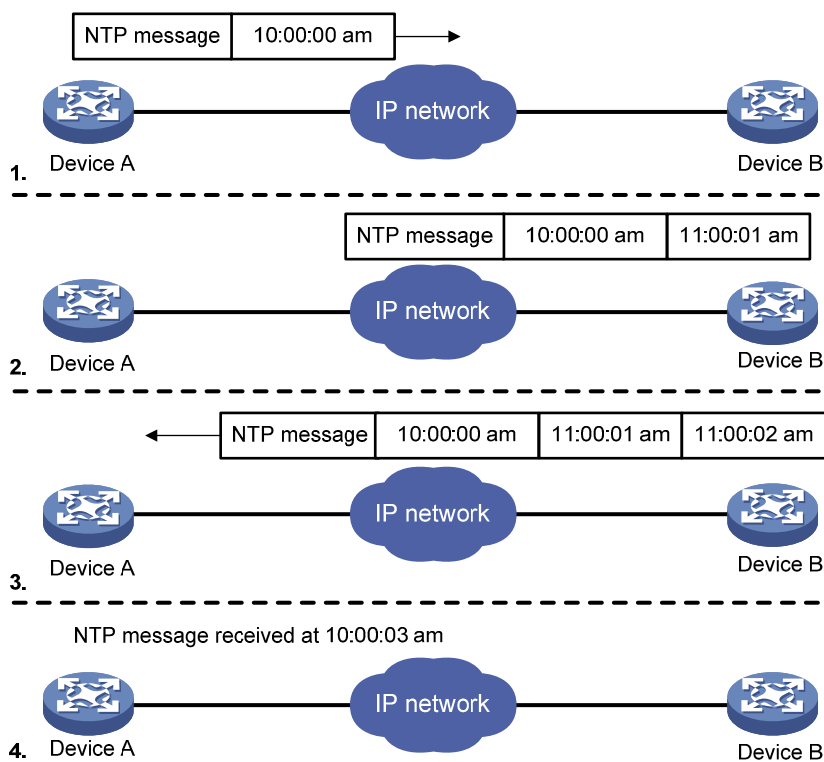
- NTP uses a stratum to describe clock accuracy. The stratum ranges from 1 to 16. Clock accuracy decreases as the stratum number increases. The stratum of a reference clock ranges from 1 to 15. A stratum 16 clock is in unsynchronized state.
- The local clock of this Switch Series cannot operate as a reference clock. It can serve as an NTP server only after it is synchronized.
- NTP supports access control and MD5 authentication.
- NTP can unicast, multicast, or broadcast protocol messages.

How NTP works

Figure 5 shows the basic workflow of NTP. Device A and Device B are connected over a network. They have their own independent system clocks, which need to be automatically synchronized through NTP. Assume that:

- Prior to system clock synchronization between Device A and Device B, the clock of Device A is set to 10:00:00 am while that of Device B is set to 11:00:00 am.
- Device B is used as the NTP time server, so Device A synchronizes to Device B.
- It takes 1 second for an NTP message to travel from one device to the other.

Figure 5 Basic workflow of NTP



The synchronization process is as follows:

- Device A sends Device B an NTP message, which is timestamped when it leaves Device A. The timestamp is 10:00:00 am (T1).
- When this NTP message arrives at Device B, it is timestamped by Device B. The timestamp is 11:00:01 am (T2).
- When the NTP message leaves Device B, Device B timestamps it. The timestamp is 11:00:02 am (T3).
- When Device A receives the NTP message, the local time of Device A is 10:00:03 am (T4).

Up to now, Device A can calculate the following parameters based on the timestamps:

- The roundtrip delay of NTP message: $\text{Delay} = (T4 - T1) - (T3 - T2) = 2 \text{ seconds}$.
- Time difference between Device A and Device B: $\text{Offset} = ((T2 - T1) + (T3 - T4)) / 2 = 1 \text{ hour}$.

Based on these parameters, Device A can synchronize its own clock to the clock of Device B.

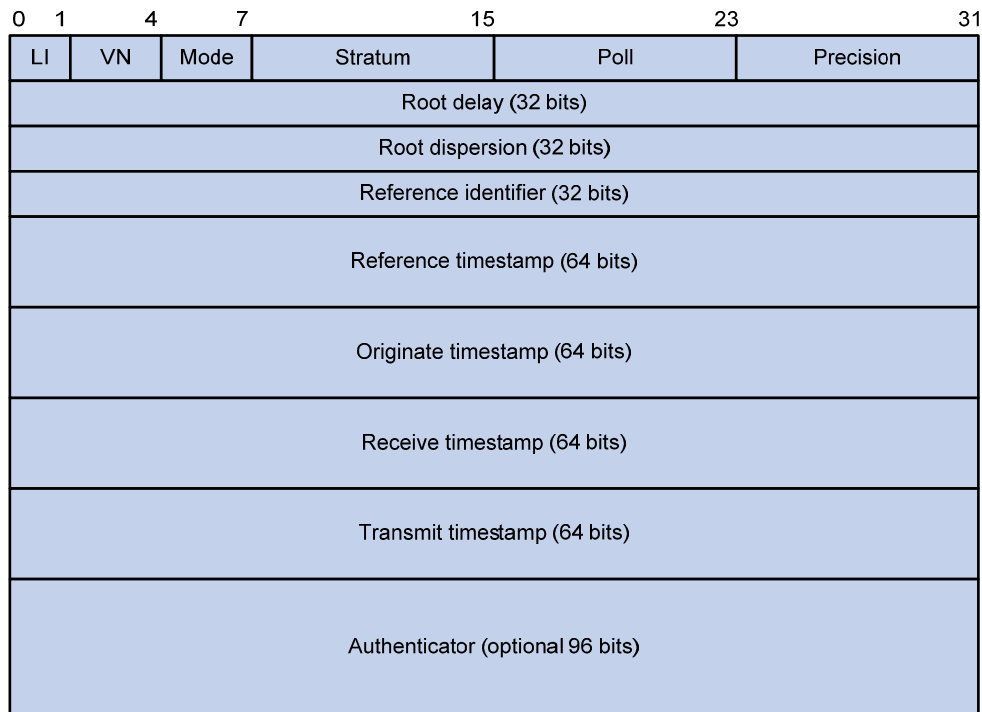
This is a rough description of how NTP works. For more information, see RFC 1305.

NTP message format

NTP uses two types of messages: clock synchronization and NTP control messages. All NTP messages mentioned in this document refer to NTP clock synchronization messages. NTP control messages are used in environments where network management is needed. Because NTP control messages are not essential for clock synchronization, they are not described in this document.

A clock synchronization message is encapsulated in a UDP message in the format shown in [Figure 6](#).

Figure 6 Clock synchronization message format



The main fields are described as follows:

- **LI (Leap Indicator)**—A 2-bit leap indicator. If set to 11, it warns of an alarm condition (clock unsynchronized). If set to any other value, it is not to be processed by NTP.
- **VN (Version Number)**—A 3-bit version number that indicates the version of NTP. The latest version is version 4.
- **Mode**—A 3-bit code that indicates the operation mode of NTP. This field can be set to these values:
 - **0**—Reserved
 - **1**—Symmetric active
 - **2**—Symmetric passive
 - **3**—Client
 - **4**—Server
 - **5**—Broadcast or multicast
 - **6**—NTP control message
 - **7**—Reserved for private use.

- **Stratum**—An 8-bit integer that indicates the stratum level of the local clock, with the value ranging from 1 to 16. Clock precision decreases from stratum 1 through stratum 16. A stratum 1 clock has the highest precision, and a stratum 16 clock is not synchronized.
- **Poll**—An 8-bit signed integer that indicates the maximum interval between successive messages, which is called the poll interval.
- **Precision**—An 8-bit signed integer that indicates the precision of the local clock.
- **Root Delay**—Roundtrip delay to the primary reference source.
- **Root Dispersion**—The maximum error of the local clock relative to the primary reference source.
- **Reference Identifier**—Identifier of the particular reference source.
- **Reference Timestamp**—The local time at which the local clock was last set or corrected.
- **Originate Timestamp**—The local time at which the request departed from the client for the service host.
- **Receive Timestamp**—The local time at which the request arrived at the service host.
- **Transmit Timestamp**—The local time at which the reply departed from the service host for the client.
- **Authenticator**—Authentication information.

Operation modes

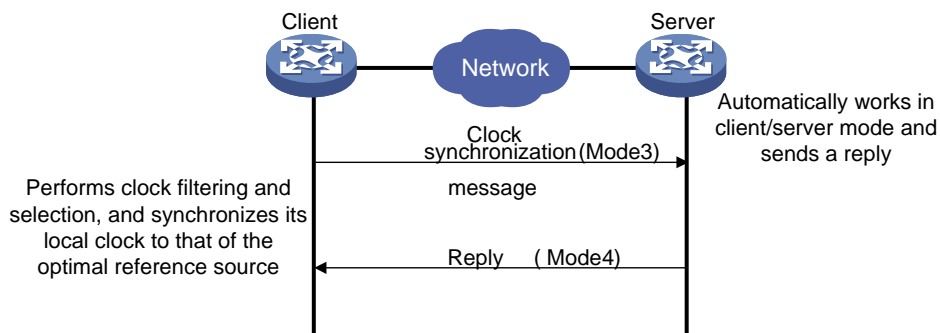
Devices that run NTP can implement clock synchronization in one of the following modes:

- Client/server mode
- Symmetric peers mode
- Broadcast mode
- Multicast mode

You can select operation modes of NTP as needed. If the IP address of the NTP server or peer is unknown and many devices in the network need to be synchronized, adopt the broadcast or multicast mode. In the client/server or symmetric peers mode, a device is synchronized from the specified server or peer, so clock reliability is enhanced.

Client/server mode

Figure 7 Client/server mode

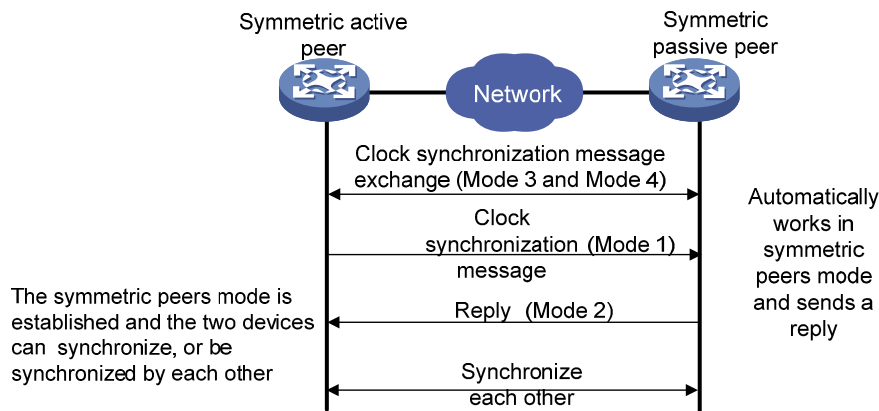


When operating in client/server mode, a client sends a clock synchronization message to servers with the Mode field in the message set to 3 (client mode). Upon receiving the message, the servers automatically operate in server mode and send a reply, with the Mode field in the messages set to 4 (server mode). Upon receiving the replies from the servers, the client performs clock filtering and selection and synchronizes to the optimal reference source.

In client/server mode, a client can synchronize to a server, but a server cannot synchronize to a client.

Symmetric peers mode

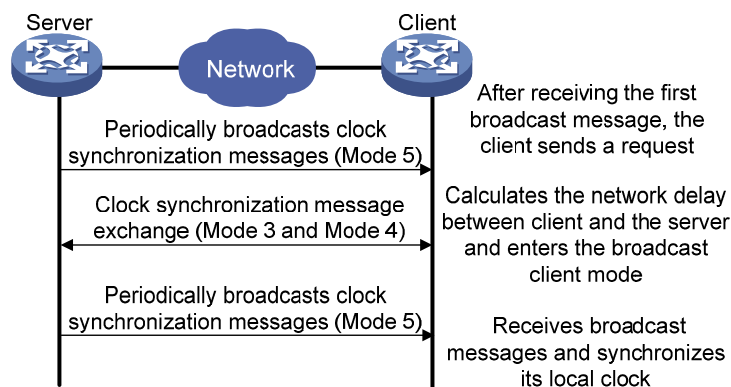
Figure 8 Symmetric peers mode



In symmetric peers mode, devices that operate in symmetric active mode and symmetric passive mode exchange NTP messages with the Mode field 3 (client mode) and 4 (server mode). Then the device that operates in symmetric active mode periodically sends clock synchronization messages, with the Mode field in the messages set to 1 (symmetric active). The device that receives the messages automatically enters symmetric passive mode and sends a reply, with the Mode field in the message set to 2 (symmetric passive). This exchange of messages establishes symmetric peers mode between the two devices, so the two devices can synchronize, or be synchronized by, each other. If the clocks of both devices have been synchronized, the device whose local clock has a lower stratum level synchronizes the other device.

Broadcast mode

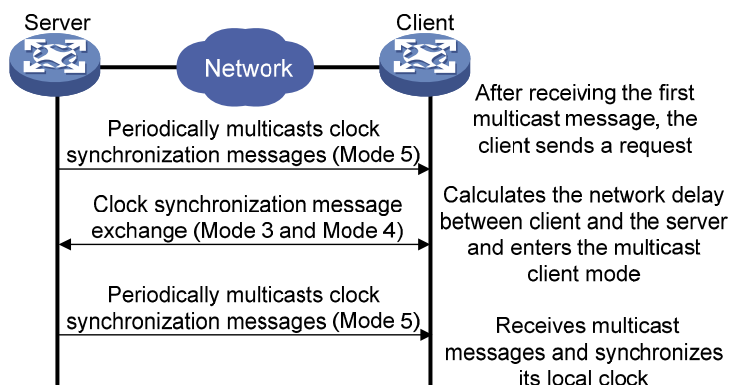
Figure 9 Broadcast mode



In broadcast mode, a server periodically sends clock synchronization messages to broadcast address 255.255.255.255, with the Mode field in the messages set to 5 (broadcast mode). Clients listen to the broadcast messages from servers. When a client receives the first broadcast message, the client and the server start to exchange messages with the Mode field set to 3 (client mode) and 4 (server mode), to calculate the network delay between client and the server. Then, the client enters broadcast client mode. The client continues listening to broadcast messages, and synchronizes its local clock based on the received broadcast messages.

Multicast mode

Figure 10 Multicast mode



In multicast mode, a server periodically sends clock synchronization messages to the user-configured multicast address, or, if no multicast address is configured, to the default NTP multicast address 224.0.1.1, with the Mode field in the messages set to 5 (multicast mode). Clients listen to the multicast messages from servers. When a client receives the first multicast message, the client and the server start to exchange messages with the Mode field set to 3 (client mode) and 4 (server mode), to calculate the network delay between client and server. Then, the client enters multicast client mode. It continues listening to multicast messages, and synchronizes its local clock based on the received multicast messages.

In symmetric peers mode, broadcast mode, and multicast mode, the client (or the symmetric active peer) and the server (the symmetric passive peer) can operate in the specified NTP operation mode only after they exchange NTP messages with the Mode field 3 (client mode) and the Mode field 4 (server mode). During this message exchange process, NTP clock synchronization can be implemented.

NTP configuration task list

| Task | Remarks |
|---|----------|
| Configuring NTP operation modes | Required |
| Configuring optional parameters | Optional |
| Configuring access-control rights | Optional |
| Configuring NTP authentication | Optional |

Configuring NTP operation modes

Devices can implement clock synchronization in one of the following modes:

- **Client/server mode**—Configure only clients.
- **Symmetric mode**—Configure only symmetric-active peers.
- **Broadcast mode**—Configure both clients and servers.
- **Multicast mode**—Configure both clients and servers.

Configuring the client/server mode

For devices operating in client/server mode, make configurations on the clients.

If you specify the source interface for NTP messages by specifying the source interface `source-interface` option, NTP uses the primary IP address of the specified interface as the source IP address of the NTP messages.

A device can act as a server to synchronize other devices only after it is synchronized. If a server has a stratum level higher than or equal to a client, the client will not synchronize to that server.

To specify an NTP server on the client:

| Step | Command | Remarks |
|--|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Specify an NTP server for the device. | ntp-service unicast-server
[vpn-instance <i>vpn-instance-name</i>]
{ <i>ip-address</i> <i>server-name</i> }
[authentication-keyid <i>keyid</i>
priority source-interface
<i>interface-type interface-number</i>
version <i>number</i>] * | By default, no NTP server is specified.

Only the HP 5500 EI supports the vpn-instance keyword.

In this command, the <i>ip-address</i> argument must be a unicast address, rather than a broadcast address, a multicast address or the IP address of the local clock.

You can configure multiple servers by repeating the command. The clients will select the optimal reference source. |

Configuring the symmetric peers mode

Follow these guidelines when you configure the NTP symmetric peers mode:

- For devices operating in symmetric mode, specify a symmetric-passive peer on a symmetric-active peer.
- Use any NTP configuration command in [Configuring NTP operation modes](#) to enable NTP. Otherwise, a symmetric-passive peer does not process NTP messages from a symmetric-active peer.
- Either the symmetric-active peer or the symmetric-passive peer must be in synchronized state. Otherwise, clock synchronization does not proceed.

To specify a symmetric-passive peer on the active peer:

| Step | Command | Remarks |
|-----------------------|--------------------|---------|
| 1. Enter system view. | system-view | N/A |

| Step | Command | Remarks |
|---|--|---|
| 2. Specify a symmetric-passive peer for the device. | <pre> ntp-service unicast-peer [vpn-instance <i>vpn-instance-name</i>] { <i>ip-address</i> <i>peer-name</i> } [authentication-keyid <i>keyid</i> priority source-interface <i>interface-type interface-number</i> version <i>number</i>] *</pre> | <p>By default, no symmetric-passive peer is specified.</p> <p>Only the HP 5500 EI supports the vpn-instance keyword.</p> <p>The <i>ip-address</i> argument must be a unicast address, rather than a broadcast address, a multicast address, or the IP address of the local clock.</p> <p>After you specify the source interface for NTP messages by specifying the source interface <i>source-interface</i> option, the source IP address of the NTP messages is set as the primary IP address of the specified interface.</p> <p>You can configure multiple symmetric-passive peers by repeating this command.</p> |

Configuring the broadcast mode

The broadcast server periodically sends NTP broadcast messages to the broadcast address 255.255.255.255. After receiving the messages, the device operating in NTP broadcast client mode sends a reply and synchronizes to the server.

Configure the NTP broadcast mode on both the server and clients. The NTP broadcast mode can only be configured in a specific interface view because an interface needs to be specified on the broadcast server for sending NTP broadcast messages and on each broadcast client for receiving broadcast messages.

Configuring a broadcast client

| Step | Command | Remarks |
|--|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter Layer 3 Ethernet interface view or VLAN interface view. | interface <i>interface-type interface-number</i> | <p>This command enters the view of the interface for sending NTP broadcast messages.</p> <p>You can configure an Ethernet port as a Layer 3 Ethernet port only on the HP 5500 EI switch.</p> |
| 3. Configure the device to operate in NTP broadcast client mode. | ntp-service broadcast-client | N/A |

Configuring the broadcast server

| Step | Command | Remarks |
|--|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter Layer 3 Ethernet interface view or VLAN interface view. | interface <i>interface-type</i>
<i>interface-number</i> | This command enters the view of the interface for sending NTP broadcast messages.

You can configure an Ethernet port as a Layer 3 Ethernet port only on the HP 5500 EI switch. |
| 3. Configure the device to operate in NTP broadcast server mode. | ntp-service broadcast-server
[authentication-keyid <i>keyid</i>
version <i>number</i>] * | A broadcast server can synchronize broadcast clients only when its clock has been synchronized. |

Configuring the multicast mode

The multicast server periodically sends NTP multicast messages to multicast clients, which send replies after receiving the messages and synchronize their local clocks.

Configure the NTP multicast mode on both the server and clients. The NTP multicast mode must be configured in a specific interface view.

Configuring a multicast client

| Step | Command | Remarks |
|--|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter Layer 3 Ethernet interface view or VLAN interface view. | interface <i>interface-type</i>
<i>interface-number</i> | This command enters the view of the interface for sending NTP multicast messages.

You can configure an Ethernet port as a Layer 3 Ethernet port only on the HP 5500 EI switch. |
| 3. Configure the device to operate in NTP multicast client mode. | ntp-service multicast-client
[<i>ip-address</i>] | You can configure up to 1024 multicast clients, of which 128 can take effect at the same time. |

Configuring the multicast server

| Step | Command | Remarks |
|--|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter Layer 3 Ethernet interface view or VLAN interface view. | interface <i>interface-type</i>
<i>interface-number</i> | This command enters the view of the interface for sending NTP multicast messages.

You can configure an Ethernet port as a Layer 3 Ethernet port only on the HP 5500 EI switch. |

| Step | Command | Remarks |
|--|--|---|
| 3. Configure the device to operate in NTP multicast server mode. | ntp-service multicast-server
[<i>ip-address</i>]
[authentication-keyid <i>keyid</i> ttl <i>ttl-number</i> version <i>number</i>] * | A multicast server can synchronize broadcast clients only when its clock has been synchronized. |

Configuring optional parameters

This section explains how to configure the optional parameters of NTP.

Specifying the source interface for NTP messages

If you specify the source interface for NTP messages, the device sets the source IP address of the NTP messages as the primary IP address of the specified interface when sending the NTP messages.

When the device responds to an NTP request received, the source IP address of the NTP response is always the IP address of the interface that received the NTP request.

Configuration guidelines

- The source interface for NTP unicast messages is the interface specified in the **ntp-service unicast-server** or **ntp-service unicast-peer** command.
- The source interface for NTP broadcast or multicast messages is the interface where you configure the **ntp-service broadcast-server** or **ntp-service multicast-server** command.
- If the specified source interface goes down, NTP uses the primary IP address of the outgoing interface as the source IP address.

Configuration procedure

To specify the source interface for NTP messages:

| Step | Command | Remarks |
|---|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Specify the source interface for NTP messages. | ntp-service source-interface
<i>interface-type interface-number</i> | By default, no source interface is specified for NTP messages, and the system uses the IP address of the interface determined by the matching route as the source IP address of NTP messages. |

Disabling an interface from receiving NTP messages

If NTP is enabled, NTP messages can be received from all interfaces by default. You can disable an interface from receiving NTP messages by using the following configuration.

| Step | Command | Remarks |
|-----------------------|--------------------|---------|
| 1. Enter system view. | system-view | N/A |

| Step | Command | Remarks |
|--|---|--|
| 2. Enter Layer 3 Ethernet interface view or VLAN interface view. | interface <i>interface-type</i>
<i>interface-number</i> | You can configure an Ethernet port as a Layer 3 Ethernet port only on the HP 5500 EI switch. |
| 3. Disable the interface from receiving NTP messages. | ntp-service in-interface disable | By default, an interface is enabled to receive NTP messages. |

Configuring the allowed maximum number of dynamic sessions

A single device can have a maximum of 128 associations at the same time, including static associations and dynamic associations.

A static association refers to an association that a user has manually created by using an NTP command. A dynamic association is a temporary association created by the system during operation. A dynamic association is removed if the system fails to receive messages from it over a specific long time.

In client/server mode, for example, when you execute a command to synchronize the time to a server, the system creates a static association, and the server simply responds passively upon the receipt of a message, rather than creating an association (static or dynamic). In symmetric mode, static associations are created at the symmetric-active peer side, and dynamic associations are created at the symmetric-passive peer side. In broadcast or multicast mode, static associations are created at the server side, and dynamic associations are created at the client side.

To configure the allowed maximum number of dynamic sessions:

| Step | Command | Remarks |
|--|--|---------------------|
| 1. Enter system view. | system-view | N/A |
| 2. Configure the maximum number of dynamic sessions allowed to be established locally. | ntp-service max-dynamic-sessions
<i>number</i> | The default is 100. |

Configuring the DSCP value for NTP messages

| Step | Command | Remarks |
|---|---|----------------------------|
| 1. Enter system view. | system-view | N/A |
| 2. Configure the Differentiated Service Code Point (DSCP) value for NTP messages. | ntp-service dscp <i>dscp-value</i> | The default setting is 16. |

Configuring access-control rights

From the highest to lowest, the NTP service access-control rights are **peer**, **server**, **synchronization**, and **query**. If a device receives an NTP request, it performs an access-control right match and uses the first matched right. If no matched right is found, the device drops the NTP request.

- **query**—Control query permitted. This level of right permits the peer devices to perform control query to the NTP service on the local device, but it does not permit a peer device to synchronize to the local device. "Control query" refers to the query of some states of the NTP service, including alarm information, authentication status, and clock source information.
- **synchronization**—Server access only. This level of right permits a peer device to synchronize to the local device, but it does not permit the peer devices to perform control query.
- **server**—Server access and query permitted. This level of right permits the peer devices to perform synchronization and control query to the local device, but it does not permit the local device to synchronize to a peer device.
- **peer**—Full access. This level of right permits the peer devices to perform synchronization and control query to the local device, and it permits the local device to synchronize to a peer device.

The access-control right mechanism provides only a minimum level of security protection for a system running NTP. A more secure method is identity authentication.

Configuration prerequisites

Before you configure the NTP service access-control right to the local device, create and configure an ACL associated with the access-control right. For more information about ACLs, see *ACL and QoS Configuration Guide*.

Configuration procedure

To configure the NTP service access-control right to the local device:

| Step | Command | Remarks |
|---|--|------------------------------|
| 1. Enter system view. | system-view | N/A |
| 2. Configure the NTP service access-control right for a peer device to access the local device. | ntp-service access { peer query server synchronization }
acl-number | The default is peer . |

Configuring NTP authentication

Enable NTP authentication for a system running NTP in a network where there is a high security demand. NTP authentication enhances network security by using client-server key authentication, which prohibits a client from synchronizing with a device that fails authentication.

To configure NTP authentication, do the following:

- Enable NTP authentication
- Configure an authentication key
- Configure the key as a trusted key
- Associate the specified key with an NTP server or a symmetric peer

These tasks are required. If any task is omitted, NTP authentication cannot function.

Configuring NTP authentication in client/server mode

Follow these instructions to configure NTP authentication in client/server mode:

- A client can synchronize to the server only when you configure all the required tasks on both the client and server.
- On the client, if NTP authentication is not enabled or no key is specified to associate with the NTP server, the client is not authenticated. No matter whether NTP authentication is enabled or not on the server, the clock synchronization between the server and client can be performed.
- On the client, if NTP authentication is enabled and a key is specified to associate with the NTP server, but the key is not a trusted key, the client does not synchronize to the server no matter whether NTP authentication is enabled or not on the server.

Configuring NTP authentication for client

| Step | Command | Remarks |
|--|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enable NTP authentication. | ntp-service authentication enable | By default, NTP authentication is disabled. |
| 3. Configure an NTP authentication key. | ntp-service authentication-keyid
<i>keyid</i> authentication-mode md5
[cipher simple] <i>value</i> | By default, no NTP authentication key is configured.
Configure the same authentication key on the client and server. |
| 4. Configure the key as a trusted key. | ntp-service reliable
authentication-keyid <i>keyid</i> | By default, the authentication key is not configured as a trusted key. |
| 5. Associate the specified key with an NTP server. | <ul style="list-style-type: none"> • Client/server mode:
ntp-service unicast-server
{ <i>ip-address</i> <i>server-name</i> }
authentication-keyid <i>keyid</i> • Symmetric peers mode:
ntp-service unicast-peer
{ <i>ip-address</i> <i>peer-name</i> }
authentication-keyid <i>keyid</i> | You can associate a non-existing key with an NTP server. To make NTP authentication effective, you must configure the key as an authentication key and specify it as a trusted key after associating the key with the NTP server. |

Configuring NTP authentication for a server

| Step | Command | Remarks |
|---|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enable NTP authentication. | ntp-service authentication enable | By default, NTP authentication is disabled. |
| 3. Configure an NTP authentication key. | ntp-service authentication-keyid
<i>keyid</i> authentication-mode md5
[cipher simple] <i>value</i> | By default, no NTP authentication key is configured.
Configure the same authentication key on the client and server. |
| 4. Configure the key as a trusted key. | ntp-service reliable
authentication-keyid <i>keyid</i> | By default, the authentication key is not configured as a trusted key. |
| 5. Enter Layer 3 Ethernet port view or VLAN interface view. | interface <i>interface-type</i>
<i>interface-number</i> | You can configure an Ethernet port as a Layer 3 Ethernet port only on the HP 5500 EI switch. |

| Step | Command | Remarks |
|--|--|--|
| 6. Associate the specified key with an NTP server. | <ul style="list-style-type: none"> Broadcast server mode:
ntp-service broadcast-server authentication-keyid <i>keyid</i> Multicast server mode:
ntp-service multicast-server authentication-keyid <i>keyid</i> | You can associate a non-existing key with an NTP server. To enable NTP authentication, you must configure the key and specify it as a trusted key after associating the key with the NTP server. |

Displaying and maintaining NTP

| Task | Command | Remarks |
|---|---|-----------------------|
| Display information about NTP service status. | display ntp-service status [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display information about NTP sessions. | display ntp-service sessions [verbose] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display the brief information about the NTP servers from the local device back to the primary reference source. | display ntp-service trace [{ begin exclude include } <i>regular-expression</i>] | Available in any view |

NTP configuration examples

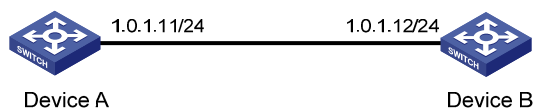
This section provides configuration examples for NTP.

Configuring the client/server mode

Network requirements

As shown in Figure 11, configure Device A as a reference source, with the stratum level 2. Configure Device B to operate in client/server mode and use Device A as its NTP server.

Figure 11 Network diagram



Configuration procedure

- Set the IP address for each interface as shown in Figure 11. (Details not shown.)
- Configure Device B:

```
# Display the NTP status of Device B before clock synchronization.
```

```
<DeviceB> display ntp-service status
Clock status: unsynchronized
Clock stratum: 16
Reference clock ID: none
```

```

Nominal frequency: 64.0000 Hz
Actual frequency: 64.0000 Hz
Clock precision: 2^7
Clock offset: 0.0000 ms
Root delay: 0.00 ms
Root dispersion: 0.00 ms
Peer dispersion: 0.00 ms
Reference time: 00:00:00.000 UTC Jan 1 1900 (00000000.00000000)
# Specify Device A as the NTP server of Device B so Device B synchronizes to Device A.

```

```

<DeviceB> system-view
[DeviceB] ntp-service unicast-server 1.0.1.11
# Display the NTP status of Device B after clock synchronization.

```

```

[DeviceB] display ntp-service status
Clock status: synchronized
Clock stratum: 3
Reference clock ID: 1.0.1.11
Nominal frequency: 64.0000 Hz
Actual frequency: 64.0000 Hz
Clock precision: 2^7
Clock offset: 0.0000 ms
Root delay: 31.00 ms
Root dispersion: 1.05 ms
Peer dispersion: 7.81 ms
Reference time: 14:53:27.371 UTC Sep 19 2005 (C6D94F67.5EF9DB22)

```

The output shows that Device B has synchronized to Device A because it has a higher stratum than Device A.

```

# Display the NTP session information of Device B.

```

```

[DeviceB] display ntp-service sessions
      source      reference  stra reach poll now offset delay disper
*****
[12345] 1.0.1.11 127.127.1.0  2   63  64   3  -75.5  31.0 16.5
note: 1 source(master),2 source(peer),3 selected,4 candidate,5 configured
Total associations : 1

```

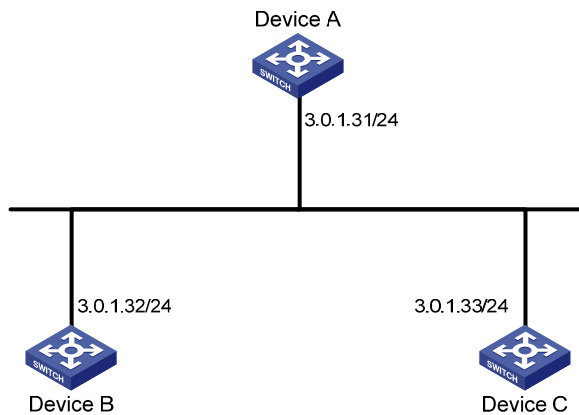
The output shows that an association has been set up between Device B and Device A.

Configuring the NTP symmetric mode

Network requirements

- As shown in [Figure 12](#), configure Device A as a reference source, with the stratum level 2.
- Configure Device B to operate in client mode and use Device A as its NTP server.
- Configure Device C to operate in symmetric-active mode and use Device B as its symmetric-passive peer.

Figure 12 Network diagram



Configuration procedure

1. Configure IP addresses for interfaces. (Details not shown.)
2. Configure Device B:

Specify Device A as the NTP server of Device B.

```
<DeviceB> system-view
[DeviceB] ntp-service unicast-server 3.0.1.31
```

3. Display the NTP status of Device B after clock synchronization.

```
[DeviceB] display ntp-service status
Clock status: synchronized
Clock stratum: 3
Reference clock ID: 3.0.1.31
Nominal frequency: 100.0000 Hz
Actual frequency: 100.0000 Hz
Clock precision: 2^18
Clock offset: -21.1982 ms
Root delay: 15.00 ms
Root dispersion: 775.15 ms
Peer dispersion: 34.29 ms
Reference time: 15:22:47.083 UTC Sep 19 2005 (C6D95647.153F7CED)
```

The output shows that Device B has synchronized to Device A because it has a higher stratum than Device A.

4. Configure Device C (after Device B is synchronized to Device A):

Configure Device C as a symmetric peer after local synchronization.

```
[DeviceC] ntp-service unicast-peer 3.0.1.32
```

The output shows that Device B and Device C are configured as symmetric peers, with Device C in symmetric-active mode and Device B in symmetric-passive mode. Because the stratus level of Device C is 16 while that of Device B is 3, Device B synchronizes to Device C.

Display the NTP status of Device C after clock synchronization.

```
[DeviceC] display ntp-service status
Clock status: synchronized
Clock stratum: 4
Reference clock ID: 3.0.1.32
```

```

Nominal frequency: 100.0000 Hz
Actual frequency: 100.0000 Hz
Clock precision: 2^18
Clock offset: -21.1982 ms
Root delay: 15.00 ms
Root dispersion: 775.15 ms
Peer dispersion: 34.29 ms
Reference time: 15:22:47.083 UTC Sep 19 2005 (C6D95647.153F7CED)

```

The output shows that Device C has synchronized to Device B because it has a higher stratum than Device B.

Display the NTP session information of Device C.

```
[DeviceC] display ntp-service sessions
```

```

          source          reference      stra reach poll  now offset  delay disper
*****
[12345] 3.0.1.32          3.0.1.31          3    3    64    16   -6.4   4.8    1.0
note: 1 source(master),2 source(peer),3 selected,4 candidate,5 configured
Total associations : 1

```

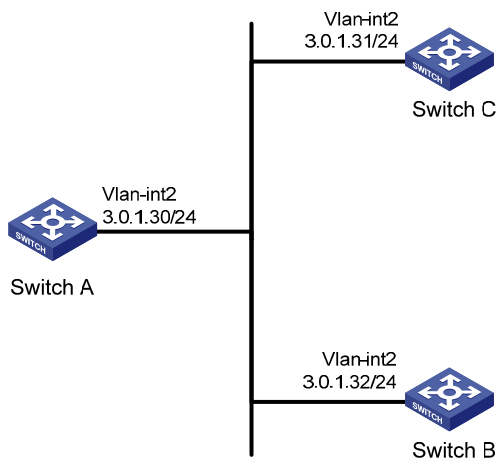
The output shows that an association has been set up between Device B and Device C.

Configuring NTP broadcast mode

Network requirements

- As shown in [Figure 13](#), configure Switch C as a reference source, with the stratum level 2.
- Configure Switch C to operate in broadcast server mode and send broadcast messages from VLAN-interface 2.
- Configure Switch A and Switch B to operate in broadcast client mode, and listen to broadcast messages through their VLAN-interface 2 respectively.

Figure 13 Network diagram



Configuration procedure

1. Set the IP address for each interface as shown in [Figure 13](#). (Details not shown.)
2. Configure Switch C:

Configure Switch C to operate in broadcast server mode and send broadcast messages through VLAN-interface 2.

```
[SwitchC] interface vlan-interface 2
[SwitchC-Vlan-interface2] ntp-service broadcast-server
```

3. Configure Switch A:

Configure Switch A to operate in broadcast client mode and receive broadcast messages on VLAN-interface 2.

```
<SwitchA> system-view
[SwitchA] interface vlan-interface 2
[SwitchA-Vlan-interface2] ntp-service broadcast-client
```

4. Configure Switch B:

Configure Switch B to operate in broadcast client mode and receive broadcast messages on VLAN-interface 2.

```
<SwitchB> system-view
[SwitchB] interface vlan-interface 2
[SwitchB-Vlan-interface2] ntp-service broadcast-client
```

Switch A and Switch B get synchronized upon receiving a broadcast message from Switch C.

Take Switch A as an example. Display the NTP status of Switch A after clock synchronization.

```
[SwitchA-Vlan-interface2] display ntp-service status
Clock status: synchronized
Clock stratum: 3
Reference clock ID: 3.0.1.31
Nominal frequency: 64.0000 Hz
Actual frequency: 64.0000 Hz
Clock precision: 2^7
Clock offset: 0.0000 ms
Root delay: 31.00 ms
Root dispersion: 8.31 ms
Peer dispersion: 34.30 ms
Reference time: 16:01:51.713 UTC Sep 19 2005 (C6D95F6F.B6872B02)
```

The output shows that Switch A has synchronized to Switch C because it has a higher stratum than Switch C.

Display the NTP session information of Switch A.

```
[SwitchA-Vlan-interface2] display ntp-service sessions
      source      reference      stra reach poll now offset delay disper
*****
[1234] 3.0.1.31 127.127.1.0 2 254 64 62 -16.0 32.0 16.6
note: 1 source(master),2 source(peer),3 selected,4 candidate,5 configured
Total associations : 1
```

The output shows that an association has been set up between Switch A and Switch C.

Configuring NTP multicast mode

Network requirements

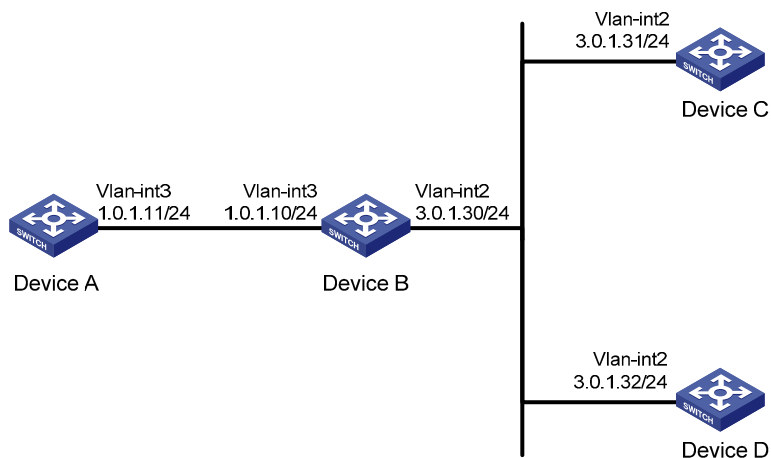
As shown in [Figure 14](#), configure Device C as a reference source, with the stratum level 2.

- Configure Device C to operate in multicast server mode and send multicast messages from VLAN-interface 2.
- Configure Device A and Device D to operate in multicast client mode and receive multicast messages through VLAN-interface 3 and VLAN-interface 2 respectively.

NOTE:

In this example, Switch B must be a Layer 3 switch that supports multicast routing.

Figure 14 Network diagram



Configuration procedure

1. Set the IP address for each interface as shown in Figure 14. (Details not shown.)
2. Configure Device C:


```
# Configure Device C to operate in multicast server mode and send multicast messages through VLAN-interface 2.
[DeviceC] interface vlan-interface 2
[DeviceC-Vlan-interface2] ntp-service multicast-server
```

3. Configure Device D:


```
# Configure Device D to operate in multicast client mode and receive multicast messages on VLAN-interface 2.
<DeviceD> system-view
[DeviceD] interface vlan-interface 2
[DeviceD-Vlan-interface2] ntp-service multicast-client
```

Because Device D and Device C are on the same subnet, Device D can receive the multicast messages from Device C without being enabled with the multicast functions and can synchronize to Device C.

Display the NTP status of Device D after clock synchronization.

```
[DeviceD-Vlan-interface2] display ntp-service status
Clock status: synchronized
Clock stratum: 3
Reference clock ID: 3.0.1.31
Nominal frequency: 64.0000 Hz
Actual frequency: 64.0000 Hz
Clock precision: 2^7
```

```

Clock offset: 0.0000 ms
Root delay: 31.00 ms
Root dispersion: 8.31 ms
Peer dispersion: 34.30 ms
Reference time: 16:01:51.713 UTC Sep 19 2005 (C6D95F6F.B6872B02)

```

The output shows that Device D has synchronized to Device C because it has a higher stratum than Device C.

Display the NTP session information of Device D.

```

[DeviceD-Vlan-interface2] display ntp-service sessions
      source      reference      stra reach poll now offset delay disper
*****
[1234] 3.0.1.31 127.127.1.0 2 254 64 62 -16.0 31.0 16.6
note: 1 source(master),2 source(peer),3 selected,4 candidate,5 configured
Total associations : 1

```

The output shows that an association has been set up between Device D and Device C.

4. Configure Device B:

Because Device A and Device C are on different subnets, you must enable the multicast functions on Device B before Device A can receive multicast messages from Device C.

Enable IP multicast routing and IGMP. For more information about how to configure IGMP and PIM, see *IP Multicast Configuration Guide*.

```

<DeviceB> system-view
[DeviceB] multicast routing-enable
[DeviceB] interface vlan-interface 2
[DeviceB-Vlan-interface2] pim dm
[DeviceB-Vlan-interface2] quit
[DeviceB] vlan 3
[DeviceB-vlan3] port gigabitethernet 1/0/1
[DeviceB-vlan3] quit
[DeviceB] interface vlan-interface 3
[DeviceB-Vlan-interface3] igmp enable
[DeviceB-Vlan-interface3] igmp static-group 224.0.1.1
[DeviceB-Vlan-interface3] quit
[DeviceB] interface gigabitethernet 1/0/1
[DeviceB-GigabitEthernet1/0/1] igmp-snooping static-group 224.0.1.1 vlan 3

```

5. Configure Device A:

```

<DeviceA> system-view
[DeviceA] interface vlan-interface 3
# Configure Device A to operate in multicast client mode and receive multicast messages on
VLAN-interface 3.

```

```

[DeviceA-Vlan-interface3] ntp-service multicast-client

```

Display the NTP status of Device A after clock synchronization.

```

[DeviceA-Vlan-interface3] display ntp-service status
Clock status: synchronized
Clock stratum: 3
Reference clock ID: 3.0.1.31
Nominal frequency: 64.0000 Hz
Actual frequency: 64.0000 Hz

```



```

Clock precision: 2^7
Clock offset: 0.0000 ms
Root delay: 40.00 ms
Root dispersion: 10.83 ms
Peer dispersion: 34.30 ms
Reference time: 16:02:49.713 UTC Sep 19 2005 (C6D95F6F.B6872B02)

```

The output shows that Device A has synchronized to Device C because it has a higher stratum than Device C.

Display the NTP session information of Device A.

```

[DeviceA-Vlan-interface3] display ntp-service sessions
      source      reference      stra reach poll now offset delay disper
*****
[1234] 3.0.1.31 127.127.1.0 2 255 64 26 -16.0 40.0 16.6
note: 1 source(master),2 source(peer),3 selected,4 candidate,5 configured
Total associations : 1

```

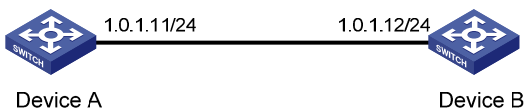
The output shows that an association has been set up between Device A and Device C.

Configuring NTP client/server mode with authentication

Network requirements

- As shown in [Figure 15](#), configure Device A as a reference source, with the stratum level 2.
- Configure Device B to operate in client mode and use Device A as its NTP server.
- Enable NTP authentication on both Device A and Device B.

Figure 15 Network diagram



Configuration procedure

1. Set the IP address for each interface as shown in [Figure 15](#). (Details not shown.)

2. Configure Device B:

```

<DeviceB> system-view
# Enable NTP authentication on Device B.
[DeviceB] ntp-service authentication enable
# Set an authentication key.
[DeviceB] ntp-service authentication-keyid 42 authentication-mode md5 aNiceKey
# Specify the key as a trusted key.
[DeviceB] ntp-service reliable authentication-keyid 42
# Specify Device A as the NTP server of Device B.
[DeviceB] ntp-service unicast-server 1.0.1.11 authentication-keyid 42

```

Before Device B can synchronize to Device A, enable NTP authentication for Device A.

3. Configure Device A:

```

# Enable NTP authentication.
[DeviceA] ntp-service authentication enable

```

```

# Set an authentication key.
[DeviceA] ntp-service authentication-keyid 42 authentication-mode md5 aNiceKey
# Specify the key as a trusted key.
[DeviceA] ntp-service reliable authentication-keyid 42
# Display the NTP status of Device B after clock synchronization.
[DeviceB] display ntp-service status
Clock status: synchronized
Clock stratum: 3
Reference clock ID: 1.0.1.11
Nominal frequency: 64.0000 Hz
Actual frequency: 64.0000 Hz
Clock precision: 2^7
Clock offset: 0.0000 ms
Root delay: 31.00 ms
Root dispersion: 1.05 ms
Peer dispersion: 7.81 ms
Reference time: 14:53:27.371 UTC Sep 19 2005 (C6D94F67.5EF9DB22)

```

The output shows that Device B has synchronized to Device A because it has a higher stratum than Device A.

```

# Display the NTP session information of Device B.
[DeviceB] display ntp-service sessions
      source      reference  stra reach poll now offset delay disper
*****
[12345] 1.0.1.11 127.127.1.0   2   63   64   3   -75.5  31.0 16.5
note: 1 source(master),2 source(peer),3 selected,4 candidate,5 configured
Total associations : 1

```

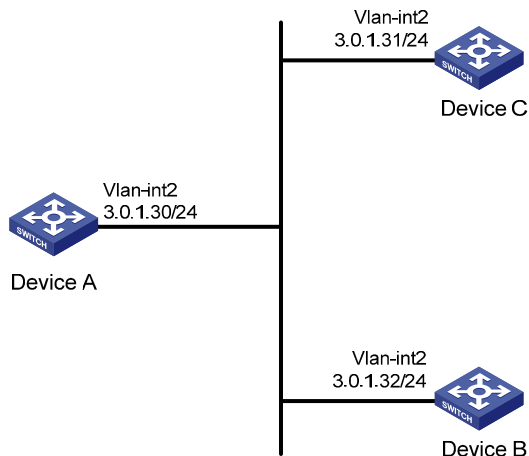
The output shows that an association has been set up Device B and Device A.

Configuring NTP broadcast mode with authentication

Network requirements

- As shown in [Figure 16](#), configure Device C as a reference source, with the stratum level 3.
- Configure Device C to operate in broadcast server mode and send broadcast messages from VLAN-interface 2.
- Configure Device A and Device B to operate in broadcast client mode and receive broadcast messages through VLAN-interface 2.
- Enable NTP authentication on both Device B and Device C.

Figure 16 Network diagram



Configuration procedure

1. Set the IP address for each interface as shown in Figure 16. (Details not shown.)
2. Configure Device A:
Configure the Device A to operate in NTP broadcast client mode and receive NTP broadcast messages on VLAN-interface 2.

```
<DeviceA> system-view
[DeviceA] interface vlan-interface 2
[DeviceA-Vlan-interface2] ntp-service broadcast-client
```
3. Configure Device B:
Enable NTP authentication on Device B. Configure an NTP authentication key, with the key ID 88 and key value 123456, and specify the key as a trusted key.

```
<DeviceB> system-view
[DeviceB] ntp-service authentication enable
[DeviceB] ntp-service authentication-keyid 88 authentication-mode md5 123456
[DeviceB] ntp-service reliable authentication-keyid 88
```


Configure Device B to operate in broadcast client mode and receive NTP broadcast messages on VLAN-interface 2.

```
[DeviceB] interface vlan-interface 2
[DeviceB-Vlan-interface2] ntp-service broadcast-client
```
4. Configure Device C:
Configure Device C to operate in NTP broadcast server mode and use VLAN-interface 2 to send NTP broadcast packets.

```
[DeviceC] interface vlan-interface 2
[DeviceC-Vlan-interface2] ntp-service broadcast-server
[DeviceC-Vlan-interface2] quit
```


Display the NTP service status information on Device A.

```
[DeviceA-Vlan-interface2] display ntp-service status
Clock status: synchronized
Clock stratum: 4
Reference clock ID: 3.0.1.31
Nominal frequency: 64.0000 Hz
```

```

Actual frequency: 64.0000 Hz
Clock precision: 2^7
Clock offset: 0.0000 ms
Root delay: 31.00 ms
Root dispersion: 8.31 ms
Peer dispersion: 34.30 ms
Reference time: 16:01:51.713 UTC Sep 19 2005 (C6D95F6F.B6872B02)

```

The output shows that Device A has synchronized to Device C because it has a higher stratum than Device C.

Display the NTP session information of Device A.

```

[DeviceA-Vlan-interface2] display ntp-service sessions
      source      reference      stra reach poll now offset delay disper
*****
[1234] 3.0.1.31 127.127.1.0 3 254 64 62 -16.0 32.0 16.6
note: 1 source(master),2 source(peer),3 selected,4 candidate,5 configured
Total associations : 1

```

The output shows that an association has been set up Device A and Device C.

.Display the NTP service status information on Device B.

```

[DeviceB-Vlan-interface2] display ntp-service status
Clock status: unsynchronized
Clock stratum: 16
Reference clock ID: none
Nominal frequency: 100.0000 Hz
Actual frequency: 100.0000 Hz
Clock precision: 2^18
Clock offset: 0.0000 ms
Root delay: 0.00 ms
Root dispersion: 0.00 ms
Peer dispersion: 0.00 ms
Reference time: 00:00:00.000 UTC Jan 1 1900(00000000.00000000)

```

The output shows that NTP authentication is enabled on Device B, but not enabled on Device C. Therefore, Device B cannot synchronize to Device C.

Enable NTP authentication Device C. Configure an NTP authentication key, with the key ID 88 and key value 123456, and specify the key as a trusted key.

```

[DeviceC] ntp-service authentication enable
[DeviceC] ntp-service authentication-keyid 88 authentication-mode md5 123456
[DeviceC] ntp-service reliable authentication-keyid 88

```

Specify Device C as an NTP broadcast server, and associate the key 88 with Device C.

```

[DeviceC] interface vlan-interface 2
[DeviceC-Vlan-interface2] ntp-service broadcast-server authentication-keyid 88

```

Display the NTP service status information on Device B.

```

[DeviceB-Vlan-interface2] display ntp-service status
Clock status: synchronized
Clock stratum: 4
Reference clock ID: 3.0.1.31
Nominal frequency: 64.0000 Hz
Actual frequency: 64.0000 Hz

```

```
Clock precision: 2^7
Clock offset: 0.0000 ms
Root delay: 31.00 ms
Root dispersion: 8.31 ms
Peer dispersion: 34.30 ms
Reference time: 16:01:51.713 UTC Sep 19 2005 (C6D95F6F.B6872B02)
```

The output shows that Device B has synchronized to Device C because it has a higher stratum than Device C.

Display the NTP session information of Device B.

```
[DeviceB-Vlan-interface2] display ntp-service sessions
      source      reference      stra reach poll now offset delay disper
*****
[1234] 3.0.1.31 127.127.1.0 3 254 64 62 -16.0 32.0 16.6
note: 1 source(master),2 source(peer),3 selected,4 candidate,5 configured
Total associations : 1
```

The output shows that an association has been set up between Device B and Device C.

Display the NTP service status information on Device A.

```
[DeviceA-Vlan-interface2] display ntp-service status
Clock status: synchronized
Clock stratum: 4
Reference clock ID: 3.0.1.31
Nominal frequency: 64.0000 Hz
Actual frequency: 64.0000 Hz
Clock precision: 2^7
Clock offset: 0.0000 ms
Root delay: 31.00 ms
Root dispersion: 8.31 ms
Peer dispersion: 34.30 ms
Reference time: 16:01:51.713 UTC Sep 19 2005 (C6D95F6F.B6872B02)
```

The output shows that configuring NTP authentication on Device C does not affect Device A and Device A still synchronizes to Device C.

Configuring the information center

This chapter describes how to configure the information center.

Overview

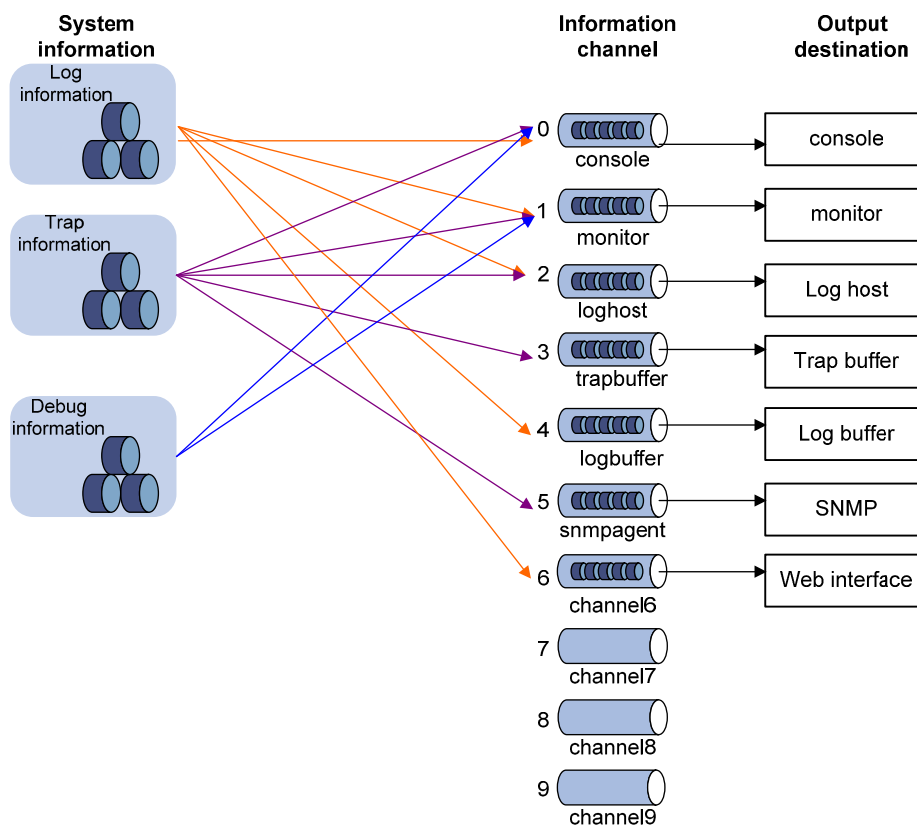
The information center classifies and manages system information so that network administrators and developers can monitor network performance and troubleshoot network problems.

The information center operates as follows:

- Receives system information including log, trap, and debugging information generated by each module.
- Outputs the information to different information channels, according to output rules.
- Outputs information to different destinations, based on information channel-to-destination associations.

Information center assigns log, trap, and debugging information to 10 information channels according to eight severity levels and then outputs the information to different destinations. The following describes the working process in detail.

Figure 17 Information center diagram (default)



By default, the information center is enabled. It affects system performance to some degree when it is processing large amounts of information. If the system resources are insufficient, disable the information center to save resources.

Classification of system information

The system information falls into the following types:

- Log information
- Trap information
- Debugging information

System information levels

The system information is classified into eight severity levels, from 0 through 7 in descending order. The device outputs the system information with a severity level that is higher than or equal to the specified level. For example, if you configure an output rule with a severity level of 6 (informational), information that has a severity level from 0 to 6 is output.

Table 1 Severity description

| Severity | Severity value | Description | Corresponding keyword in commands |
|---------------|----------------|-----------------------------------|-----------------------------------|
| Emergency | 0 | The system is unavailable. | emergencies |
| Alert | 1 | Action must be taken immediately. | alerts |
| Critical | 2 | Critical condition. | critical |
| Error | 3 | Error condition. | errors |
| Warning | 4 | Warning condition. | warnings |
| Notice | 5 | Normal but significant condition. | notifications |
| Informational | 6 | Informational messages. | informational |
| Debug | 7 | Debug messages. | debugging |

System information channels and output destinations

Table 2 shows the information channels and output destinations.

The system supports ten channels. By default, channels 0 through 6 are configured with channel names and output destinations, and they are associated with output destinations. You can change these default settings as needed. You can also configure channels 7, 8 and 9 associate them with specific output destinations to meet your needs.

Table 2 Information channels and output destinations

| Information channel number | Default channel name | Default output destination | Description |
|----------------------------|----------------------|----------------------------|---|
| 0 | console | Console | Receives log, trap and debugging information. |

| Information channel number | Default channel name | Default output destination | Description |
|----------------------------|----------------------|----------------------------|--|
| 1 | monitor | Monitor terminal | Receives log, trap and debugging information, facilitating remote maintenance. |
| 2 | loghost | Log host | Receives log, trap and debugging information and information will be stored in files for future retrieval. |
| 3 | trapbuffer | Trap buffer | Receives trap information, a buffer inside the device for recording information. |
| 4 | logbuffer | Log buffer | Receives log and debugging information, a buffer inside the device for recording information. |
| 5 | snmpagent | SNMP module | Receives trap information. |
| 6 | channel6 | Web interface | Receives log information. |
| 7 | channel7 | Not specified | Receives log, trap, and debugging information. |
| 8 | channel8 | Not specified | Receives log, trap, and debugging information. |
| 9 | channel9 | Not specified | Receives log, trap, and debugging information. |

Outputting system information by source module

The system is composed of a variety of protocol modules, and configuration modules. The system information is classified, filtered, and output according to source modules. You can use the **info-center source ?** command to view the supported information source modules.

Default output rules of system information

A default output rule specifies the system information source modules, information type, and severity levels for an output destination. [Table 3](#) shows the default output rules.

The following applies to all modules by default:

- All log information is allowed to be output to the Web interface. Log information that has a severity level of at least informational is allowed to be output to the log host. Log information that has a severity level of at least informational is allowed to be output to the console, monitor terminal, and log buffer. Log information is not allowed to be output to the trap buffer or the SNMP module.
- All trap information is allowed to be output to the console, monitor terminal, log host, Web interface. Trap information that has a severity level of at least informational is allowed to be output to the trap buffer and SNMP module. Trap information is not allowed to be output to the log buffer.
- All debugging information is allowed to be output to the console and monitor terminal. Debugging information is not allowed to be output to the log host, trap buffer, log buffer, the SNMP module, Web interface.

Table 3 Default output rules for different output destinations

| Output destination | Modules allowed | LOG | | TRAP | | DEBUG | |
|--------------------|-----------------------|------------------|---------------|------------------|---------------|------------------|----------|
| | | Enabled/disabled | Severity | Enabled/disabled | Severity | Enabled/disabled | Severity |
| Console | default (all modules) | Enabled | Informational | Enabled | Debug | Enabled | Debug |
| Monitor terminal | default (all modules) | Enabled | Informational | Enabled | Debug | Enabled | Debug |
| Log host | default (all modules) | Enabled | Informational | Enabled | Debug | Disabled | Debug |
| Trap buffer | default (all modules) | Disabled | Informational | Enabled | Informational | Disabled | Debug |
| Log buffer | default (all modules) | Enabled | Informational | Disabled | Debug | Disabled | Debug |
| SNMP module | default (all modules) | Disabled | Debug | Enabled | Informational | Disabled | Debug |
| Web interface | default (all modules) | Enabled | Debug | Enabled | Debug | Disabled | Debug |

System information format

The following shows the original format of system information, which might be different from what you see. The actual format depends on the log resolution tool you use.

Formats

The system information format depends on the output destinations, as shown in [Table 4](#).

Table 4 System information formats

| Output destination | Format | Example |
|---|---|--|
| Console, monitor terminal, logbuffer, trapbuffer, SNMP module | timestamp sysname
module/level/digest: content | %Jun 26 17:08:35:809 2011 Sysname
SHELL/4/LOGIN: VTY login from 1.1.1.1 |

| Output destination | Format | Example |
|--------------------|---|--|
| Log host | <ul style="list-style-type: none"> HP format:
<PRI>timestamp
Sysname %%vmodule/level
/digest: source content UNICOM format:
<PRI>timestamp Sysname
vmodule/level/serial_number: content | <ul style="list-style-type: none"> HP format:
<189>Oct 9 14:59:04 201
MyDevice %%10SHELL/5/SHELL_LOGIN():
VTY logged in from 192.168.1.21 UNICOM format: <ul style="list-style-type: none"> <186>Oct 13 16:48:08 2011 HP
10IFNET/2/210231a64jx073000020:
log_type=port;content=Vlan-interface 1
link status is DOWN. <186>Oct 13 16:48:08 2011 HP
10IFNET/2/210231a64jx073000020:
log_type=port;content=Line protocol on
the interface Vlan-interface1 is DOWN. |

NOTE:

The closing set of angle brackets (< >), the space, the forward slash (/), and the colon (:) are all required in the above format.

Following is a detailed explanation of the fields involved in system information:

PRI (priority)

The priority is calculated using the formula: $facility * 8 + level$, where:

- facility** is the facility name, ranging from local0 to local7 (16 to 23 in decimal integers) and defaults to local7. It can be configured with **info-center loghost**. It is used to identify different log sources on the log host, and to query and filter logs from specific log sources.
- level** ranges from 0 to 7. See [Table 1](#) for more information.

Note that the priority field is available only for information that is sent to the log host.

timestamp

Timestamp records the time when the system information was generated. The time stamp of the system information sent to the log host has a precision of seconds, and that to all the other destinations has a precision of milliseconds. The time stamp format of the system information sent to the log host is configured with the **info-center timestamp loghost** command, and that of the system information sent to the other destinations is configured with the **info-center timestamp** command.

Table 5 Description on the time stamp parameters

| Time stamp parameter | Description | Example |
|----------------------|---|--|
| boot | Time since system startup, in the format of xxxxxx.yyyyyy, where xxxxxx represents the higher 32 bits, and yyyyyy represents the lower 32 bits.
System information that is sent to all destinations except the log host supports this parameter. | %0.109391473 Sysname FTPD/5/FTPD_LOGIN: User ftp (192.168.1.23) has logged in successfully.
0.109391473 is a time stamp in the boot format. |
| date | Current date and time of the system, in the format of Mmm dd hh:mm:ss:sss yyyy.
All system information supports this parameter. | %May 30 05:36:29:579 2011 Sysname FTPD/5/FTPD_LOGIN: User ftp (192.168.1.23) has logged in successfully.
May 30 05:36:29:579 2011 is a time stamp in the date format. |
| iso | Time stamp format stipulated in ISO 8601
All system information supports this parameter. | <189>2011-05-30T06:42:44
Sysname %%10FTPD/5/FTPD_LOGIN(): User ftp (192.168.1.23) has logged in successfully.
2011-05-30T06:42:44 is a time stamp in the iso format. |
| none | No time stamp is included.
All system information supports this parameter. | % Sysname FTPD/5/FTPD_LOGIN: User ftp (192.168.1.23) has logged in successfully.
No time stamp is included. |
| no-year-date | Current date and time of the system, with year information excluded.
Only the system information that is sent to the log host supports this parameter. | <189>May 30 06:44:22
Sysname %%10FTPD/5/FTPD_LOGIN(): User ftp (192.168.1.23) has logged in successfully.
May 30 06:44:22 is a time stamp in the no-year-date format. |

Sysname (host name or host IP address)

- If the system information that is sent to a log host is in the UNICOM format, and the **info-center loghost source** command is configured, or **vpn-instance vpn-instance-name** is provided in the **info-center loghost** command, the field is displayed as the IP address of the device that generates the system information.
- If the system information is sent to other destinations, or is sent to a log host in the HP format, the field is displayed as the system name of the device that generates the system information. You can use the **sysname** command to modify the system name. For more information, see *Fundamentals Command Reference*.

%% (vendor ID)

This field indicates that the information is generated by an HP device. It is displayed only when the system information is sent to a log host in the format of HP.

vv

This field is a version identifier of syslog, with a value of 10. It is displayed only when the output destination is log host.

module

The module field represents the name of the module that generates system information. You can enter the **info-center source ?** command in system view to view the module list.

level (severity)

System information is divided into eight levels based on its severity, from 0 to 7. See [Table 1](#) for definitions and descriptions of these severity levels. The levels of system information generated by modules are predefined by developers, and you cannot change the system information levels. However, with the **info-center source** command, you can configure to output information of the specified level and not to output information lower than the specified level.

digest

The digest field is a string of up to 32 characters, outlining the system information.

For system information destined to the log host:

- If the character string ends with (l), the information is log information.
- If the character string ends with (t), the information is trap information.
- If the character string ends with (d), the information is debugging information.

For system information destined to other destinations:

- If the time stamp starts with a percent sign (%), the information is log information.
- If the time stamp starts with a pound sign (#), the information is trap information.
- If the time stamp starts with an asterisk (*), the information is debugging information.

serial number

This field indicates the serial number of the device that generates the system information. It is displayed only when the system information is sent to a log host in the UNICOM format.

source

This field indicates the source of the information. It is optional and is displayed only when the system information is sent to a log host in the HP format. This field takes one of the following values:

- IRF member ID
- IP address of the log sender

content

This field provides the content of the system information.

Information center configuration task list

| Task | Remarks |
|---|----------|
| Outputting system information to the console | Optional |
| Outputting system information to the monitor terminal | Optional |
| Outputting system information to a log host | Optional |

| Task | Remarks |
|---|----------|
| Outputting system information to the trap buffer | Optional |
| Outputting system information to the log buffer | Optional |
| Outputting system information to the SNMP module | Optional |
| Outputting system information to the Web interface | Optional |
| Saving security logs into the security log file | Optional |
| Configuring synchronous information output | Optional |
| Disabling an interface from generating link up/down logging information | Optional |

Outputting system information to the console

This section describes how to output system information to the console.

Configuring a system information output rule for the console

| Step | Command | Remarks |
|---|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enable information center. | info-center enable | Optional.
Enabled by default. |
| 3. Name the channel with a specified channel number. | info-center channel <i>channel-number</i>
name <i>channel-name</i> | Optional.
See Table 2 for default channel names. |
| 4. Configure the channel through which system information can be output to the console. | info-center console channel
{ <i>channel-number</i> <i>channel-name</i> } | Optional.
By default, system information is output to the console through channel 0 (known as console). |
| 5. Configure an output rule for the console. | info-center source { <i>module-name</i> default } channel { <i>channel-number</i> <i>channel-name</i> } [debug { level <i>severity</i> state <i>state</i> } * log { level <i>severity</i> state <i>state</i> } * trap { level <i>severity</i> state <i>state</i> } *] * | Optional.
See " Default output rules of system information. " |
| 6. Configure the timestamp format. | info-center timestamp { debugging log trap } { boot date none } | Optional.
By default, the timestamp format for log, trap and debugging information is date . |

Enabling system information output to the console

To enable the display of system information on the console in user view:

| Step | Command | Remarks |
|---|---|--|
| 1. Enable system information output to the console. | terminal monitor | Optional.
The default setting is enabled. |
| 2. Enable the display of system information on the console. | <ul style="list-style-type: none"> Enable the display of debugging information on the console:
terminal debugging Enable the display of log information on the console:
terminal logging Enable the display of trap information on the console:
terminal trapping | Optional.
By default, the console only displays log and trap information. |

Outputting system information to the monitor terminal

Monitor terminals refer to terminals that log in to the switch through the VTY user interface.

Configuring a system information output rule for the monitor terminal

| Step | Command | Remarks |
|--|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enable information center. | info-center enable | Optional.
Enabled by default. |
| 3. Name the channel with a specified channel number. | info-center channel name <i>channel-number</i> <i>channel-name</i> | Optional.
See Table 2 for default channel names. |
| 4. Configure the channel through which system information can be output to a monitor terminal. | info-center monitor channel { <i>channel-number</i> <i>channel-name</i> } | Optional.
By default, system information is output to the monitor terminal through channel 1 (known as monitor). |
| 5. Configure a system information output rule for the monitor terminal. | info-center source { <i>module-name</i> default } channel { <i>channel-number</i> <i>channel-name</i> } [debug { level <i>severity</i> state <i>state</i> } * log { level <i>severity</i> state <i>state</i> } * trap { level <i>severity</i> state <i>state</i> } *] * | Optional.
See " Default output rules of system information. " |
| 6. Configure the timestamp format. | info-center timestamp { debugging log trap } { boot date none } | Optional.
By default, the time stamp format for log, trap and debugging information is date . |

Enabling system information output to the monitor terminal

| Step | Command | Remarks |
|---|--|--|
| 1. Enable the monitoring of system information on a monitor terminal. | terminal monitor | Enabled on the console and disabled on the monitor terminal by default. |
| 2. Enable the display of system information on the monitor terminal. | <ul style="list-style-type: none"> Enable the display of debugging information on a monitor terminal:
terminal debugging Enable the display of log information on a monitor terminal:
terminal logging Enable the display of trap information on a monitor terminal:
terminal trapping | <p>Optional.</p> <p>By default, the monitor terminal displays only the log and trap information.</p> |

Outputting system information to a log host

| Step | Command | Remarks |
|--|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enable information center. | info-center enable | Optional.
Enabled by default. |
| 3. Name the channel with a specified channel number. | info-center channel <i>channel-number</i>
name <i>channel-name</i> | Optional.
See Table 2 for default channel names. |
| 4. Configure a system information output rule for the log host. | info-center source { <i>module-name</i> default } channel { <i>channel-number</i> <i>channel-name</i> } [debug { level <i>severity</i> state <i>state</i> } * log { level <i>severity</i> state <i>state</i> } * trap { level <i>severity</i> state <i>state</i> } *] * | Optional.
See " Default output rules of system information. " |
| 5. Specify the source IP address for the log information. | info-center loghost source
<i>interface-type interface-number</i> | Optional.
By default, the source interface is determined by the matched route, and the primary IP address of this interface is the source IP address of the log information. |
| 6. Configure the format of the time stamp for system information output to the log host. | info-center timestamp loghost { date iso no-year-date none } | Optional.
date by default. |
| 7. Set the format of the system information sent to a log host to UNICOM. | info-center format unicom | Optional.
HP by default. |

| Step | Command | Remarks |
|--|---|---|
| 8. Specify a log host and configure the related output parameters. | info-center loghost [vpn-instance <i>vpn-instance-name</i>]
{ <i>host-ipv4-address</i> ipv6 <i>host-ipv6-address</i> } [port <i>port-number</i>] [dscp <i>dscp-value</i>]
[channel { <i>channel-number</i> <i>channel-name</i> }] facility <i>local-number</i>] * | By default, the system does not output information to a log host. If you specify to output system information to a log host, the system uses channel 2 (loghost) by default.

The value of the <i>port-number</i> argument should be the same as the value configured on the log host, otherwise, the log host cannot receive system information.

The vpn-instance keyword is available only on the HP 5500 EI switch series. |

Outputting system information to the trap buffer

The trap buffer only receives trap information, and discards log and debug information.

To output system information to the trap buffer:

| Step | Command | Remarks |
|---|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enable information center. | info-center enable | Optional.
Enabled by default. |
| 3. Name the channel with a specified channel number. | info-center channel <i>channel-number</i>
name <i>channel-name</i> | Optional.
See Table 2 for default channel names. |
| 4. Configure the channel through which system information can be output to the trap buffer and specify the buffer size. | info-center trapbuffer [channel { <i>channel-number</i> <i>channel-name</i> }] size <i>buffersize</i>] * | Optional.
By default, system information is output to the trap buffer through channel 3 (known as trapbuffer) and the default buffer size is 256. |
| 5. Configure a system information output rule for the trap buffer. | info-center source { <i>module-name</i> default } channel { <i>channel-number</i> <i>channel-name</i> } [debug { level <i>severity</i> state <i>state</i> } *] log { level <i>severity</i> state <i>state</i> } * trap { level <i>severity</i> state <i>state</i> } *] * | Optional.
See " Default output rules of system information. " |
| 6. Configure the timestamp format. | info-center timestamp { debugging log trap } { boot date none } | Optional.
The time stamp format for log, trap and debugging information is date by default. |

Outputting system information to the log buffer

The log buffer only receives log information, and discards trap and debug information.

To output system information to the log buffer:

| Step | Command | Remarks |
|--|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enable information center. | info-center enable | Optional.
Enabled by default. |
| 3. Name the channel with a specified channel number. | info-center channel <i>channel-number</i>
name <i>channel-name</i> | Optional.
See Table 2 for default channel names. |
| 4. Configure the channel through which system information can be output to the log buffer and specify the buffer size. | info-center logbuffer [channel { <i>channel-number</i> <i>channel-name</i> } size <i>buffersize</i>] * | Optional.
By default, system information is output to the log buffer through channel 4 (known as logbuffer) and the default buffer size is 512. |
| 5. Configure a system formation output rule for the log buffer. | info-center source { <i>module-name</i> default } channel { <i>channel-number</i> <i>channel-name</i> } [debug { level <i>severity</i> state <i>state</i> } * log { level <i>severity</i> state <i>state</i> } * trap { level <i>severity</i> state <i>state</i> } *] * | Optional.
See " Default output rules of system information. " |
| 6. Configure the timestamp format. | info-center timestamp { debugging log trap } { boot date none } | Optional.
The time stamp format for log, trap and debugging information is date by default. |

Outputting system information to the SNMP module

The SNMP module only receives trap information, and discards log and debug information.

To monitor the device running status, trap information is usually sent to the SNMP network management system (NMS). For this purpose, you must configure output of traps to the SNMP module, and set the trap sending parameters for the SNMP module. For more information about SNMP, see "[Configuring SNMP.](#)"

To output system information to the SNMP module:

| Step | Command | Remarks |
|---|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enable information center. | info-center enable | Optional.
Enabled by default. |
| 3. Name the channel with a specified channel number. | info-center channel <i>channel-number</i>
name <i>channel-name</i> | Optional.
See Table 2 for default channel names. |
| 4. Configure the channel through which system information can be output to the SNMP module. | info-center snmp channel { <i>channel-number</i> <i>channel-name</i> } | Optional.
By default, system information is output to the SNMP module through channel 5 (known as snmpagent). |

| Step | Command | Remarks |
|--|---|---|
| 5. Configure a system formation output rule for the SNMP module. | info-center source { <i>module-name</i> default } channel { <i>channel-number</i> <i>channel-name</i> } [debug { level <i>severity</i> state <i>state</i> } * log { level <i>severity</i> state <i>state</i> } * trap { level <i>severity</i> state <i>state</i> } *] * | Optional.
See "Default output rules of system information." |
| 6. Configure the timestamp format. | info-center timestamp { debugging log trap } { boot date none } | Optional.
The time stamp format for log, trap and debugging information is date by default. |

Outputting system information to the Web interface

The Web interface only receives log information, and discards trap and debug information.

This feature allows you to control whether to output system information to the Web interface and, if so, which system information can be output to the Web interface. The Web interface provides search and sorting functions. You can view system information by clicking corresponding tabs after logging in to the device through the Web interface.

To output system information to the Web interface:

| Step | Command | Remarks |
|---|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enable information center. | info-center enable | Optional.
Enabled by default. |
| 3. Name the channel with a specified channel number. | info-center channel <i>channel-number</i>
name <i>channel-name</i> | Optional.
See Table 2 for default channel names. |
| 4. Configure the channel through which system information can be output to the Web interface. | info-center syslog channel
{ <i>channel-number</i> <i>channel-name</i> } | Optional.
By default, system information is output to the Web interface through channel 6. |
| 5. Configure a system formation output rule for the Web interface. | info-center source { <i>module-name</i> default } channel { <i>channel-number</i> <i>channel-name</i> } [debug { level <i>severity</i> state <i>state</i> } * log { level <i>severity</i> state <i>state</i> } * trap { level <i>severity</i> state <i>state</i> } *] * | Optional.
See "Default output rules of system information." |
| 6. Configure the format of the time stamp. | info-center timestamp { debugging log trap } { boot date none } | Optional.
The time stamp format for log, trap and debugging information is date by default. |

Saving security logs into the security log file

Security logs are very important for locating and troubleshooting network problems. Generally, security logs are output together with other logs. It is difficult to identify security logs among all logs.

To solve this problem, you can save security logs into a security log file without affecting the current log output rules.

The configuration of this feature and the management of the security log file are separate, and the security log file is managed by a privileged user. After logging in to the device, the administrator can enable the saving security logs into the security log file and configure related parameters. However, only the privileged user, known as the security log administrator, can perform operations on the security log file. The privileged user must pass AAA local authentication and log in to the device. No other users (including the system administrator) can perform operations on the security log file.

A security log administrator is a local user who is authorized by AAA to play the security log administrator role. You can authorize a security log administrator by executing the **authorization-attribute user-role security-audit** command in local user view.

The system administrator cannot view, copy, and rename the security log file. If they try, the system displays an "% Execution error" message. The system administrator can view, copy and rename other types of files.

For more information about local user and AAA local authentication, see *Security Configuration Guide*.

Saving security logs into the security log file

If this feature is enabled, the system first outputs security logs to the security log file buffer, and then saves the logs in the security log file buffer into the security log file at a specified interval (the security log administrator can also manually save security logs into the log file). After the logs are saved, the buffer is cleared immediately.

The size of the security log file is limited. When the maximum size is reached, the system deletes the oldest log and writes the new log into the security log file. To avoid security log loss, you can set an alarm threshold for the security log file usage. When the alarm threshold is reached, the system outputs a message to inform the administrator. The administrator can log in to the device as the security log administrator and back up the security log file to prevent the loss of important data.

By default, security logs are not saved into the security log file. The parameters, such as the saving interval, the maximum size, and the alarm threshold, have default settings. To modify these parameters, log in to the device as the system administrator, and then follow the steps in the following table to configure the related parameters:

To save security logs into the security log file:

| Step | Command | Remarks |
|---|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enable the information center. | info-center enable | Optional.
Enabled by default. |
| 3. Enable the saving of the security logs into the security log file. | info-center security-logfile enable | Disabled by default. |
| 4. Set the frequency with which the system saves the security log file. | info-center security-logfile frequency freq-sec | Optional.
The default value is 600 seconds. |

| Step | Command | Remarks |
|--|--|--|
| 5. Set the maximum storage space reserved for the security log file. | info-center security-logfile size-quota <i>size</i> | Optional.
The default value is 1 MB. |
| 6. Set the alarm threshold of the security log file usage. | info-center security-logfile alarm-threshold <i>usage</i> | Optional.
80 by default. (That is, when the usage of the security log file reaches 80%, the system will inform the user.) |

Managing the security log file

After passing the AAA local authentication, the security log administrator can perform the following operations:

| Task | Command | Remarks |
|---|--|--|
| Display a summary of the security log file. | display security-logfile summary [{ begin exclude include } <i>regular-expression</i>] | Optional. |
| Change the directory where the security log file is saved. | info-center security-logfile switch-directory <i>dir-name</i> | Optional.
By default, the directory to save the security log file is the seclog directory in the root directory of the storage medium.
Available in user view. |
| Display contents of the security log file buffer. | display security-logfile buffer [{ begin exclude include } <i>regular-expression</i>] | Optional. |
| Save all the contents in the security log file buffer into the security log file. | security-logfile save | Optional.
By default, the system automatically saves the security log file at a frequency configured by the info-center security-logfile frequency command into a directory configured by the info-center security-logfile switch-directory command.
Available in user view. |

| Task | Command | Remarks |
|---|---|--|
| Perform these operations on the security log file. | <ul style="list-style-type: none"> • Display the contents of the specified file:
more <i>file-url</i> • Display information about all files and folders:
dir [/all] [<i>file-url</i>] • Create a folder under a specified directory on the storage medium:
mkdir <i>directory</i> • Change the current working directory:
cd { <i>directory</i> .. / } • Display the current path:
pwd • Move a specified file from a storage medium to the recycle bin:
delete [/unreserved] <i>file-url</i> • Remove a folder:
rmdir <i>directory</i> • Format a storage medium:
format <i>device</i> • Restore a file from the Recycle Bin:
undelete <i>file-url</i> | <p>Optional.</p> <p>Available in user view</p> <p>For more information about these commands, see <i>Fundamentals Command Reference</i>.</p> |
| Uploading the security log file to the SFTP server. | <ul style="list-style-type: none"> • Establish an SFTP connection in an IPv4 network:
sftp <i>server</i> [<i>port-number</i>] [vpn-instance <i>vpn-instance-name</i>] [identity-key { <i>dsa</i> <i>rsa</i> } prefer-ctos-cipher { <i>3des</i> <i>aes 128</i> <i>des</i> } prefer-ctos-hmac { <i>md5</i> <i>md5-96</i> <i>sha1</i> <i>sha1-96</i> } prefer-kex { <i>dh-group-exchange</i> <i>dh-group1</i> <i>dh-group14</i> } prefer-stoc-cipher { <i>3des</i> <i>aes 128</i> <i>des</i> } prefer-stoc-hmac { <i>md5</i> <i>md5-96</i> <i>sha1</i> <i>sha1-96</i> }] * • Establish an SFTP connection in an IPv6 network:
sftp <i>server</i> [<i>port-number</i>] [vpn-instance <i>vpn-instance-name</i>] [identity-key { <i>dsa</i> <i>rsa</i> } prefer-ctos-cipher { <i>3des</i> <i>aes 128</i> <i>des</i> } prefer-ctos-hmac { <i>md5</i> <i>md5-96</i> <i>sha1</i> <i>sha1-96</i> } prefer-kex { <i>dh-group-exchange</i> <i>dh-group1</i> <i>dh-group14</i> } prefer-stoc-cipher { <i>3des</i> <i>aes 128</i> <i>des</i> } prefer-stoc-hmac { <i>md5</i> <i>md5-96</i> <i>sha1</i> <i>sha1-96</i> }] * • Upload a file on the client to the remote SFTP server:
put <i>localfile</i> [<i>remotefile</i>] • Download a file from a remote SFTP server and save it:
get <i>remotefile</i> [<i>localfile</i>] • For all other operations supported by the device acting as an SFTP client, see <i>Security Configuration Guide</i>. | <p>Optional.</p> <p>The sftp commands are available in user view; the other commands are available in SFTP client view.</p> <p>For more information about these commands, see <i>Security Command Reference</i>.</p> <p>The vpn-instance keyword is available only on the HP 5500 EI switch.</p> |

Configuring synchronous information output

The output of system logs interrupts ongoing configuration operations, and you have to find the previously input commands before the logs. Synchronous information output can show the previous input after log output and a command prompt in command editing mode, or a [Y/N] string in interaction mode so you can continue your operation from where you were stopped.

To enable synchronous information output:

| Step | Command | Remarks |
|---|--------------------------------|---------------------|
| 1. Enter system view. | system-view | N/A |
| 2. Enable synchronous information output. | info-center synchronous | Disabled by default |

If system information, such as log information, is output before you input any information under the current command line prompt, the system does not display the command line prompt.

If system information is output when you are inputting some interactive information (non Y/N confirmation information), the system displays your previous input in a new line but does not display the command line prompt.

Disabling an interface from generating link up/down logging information

By default, all interfaces generate link up or link down log information when the state changes. In some cases, you might want to disable specific interfaces from generating this information. For example:

- You are concerned only about the states of some interfaces. In this case, you can use this function to disable other interfaces from generating link up and link down log information.
- An interface is unstable and continuously outputs log information. In this case, you can disable the interface from generating link up and link down log information.

To disable an interface from generating link up/down logging information:

| Step | Command | Remarks |
|---|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter Layer 3 Ethernet interface view or Layer 2 Ethernet interface view or VLAN interface view. | interface <i>interface-type</i>
<i>interface-number</i> | You can configure an Ethernet interface as a Layer 3 Ethernet interface only on the HP 5500 EI switch. |
| 3. Disable the interface from generating link up or link down logging information. | undo enable log updown | By default, all interfaces generate link up and link down logging information when the state changes. |

Use the default setting in normal cases to avoid affecting interface status monitoring.

Displaying and maintaining information center

| Task | Command | Remarks |
|---|--|------------------------|
| Display the information about information channels. | display channel [<i>channel-number</i> <i>channel-name</i>] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display the information about each output destination. | display info-center [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display the state of the log buffer and the log information recorded. | display logbuffer [reverse] [level <i>severity</i> size <i>buffersize</i> slot <i>slot-number</i>] * [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display a summary of the log buffer. | display logbuffer summary [level <i>severity</i> slot <i>slot-number</i>] * [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display the state of the trap buffer and the trap information recorded. | display trapbuffer [reverse] [size <i>buffersize</i>] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Reset the log buffer. | reset logbuffer | Available in user view |
| Reset the trap buffer. | reset trapbuffer | Available in user view |

Information center configuration examples

Outputting log information to a UNIX log host

Network requirements

Configure the device to send ARP and IP log information that has a severity level of at least informational to the UNIX log host at 1.2.0.1/16.

Figure 18 Network diagram



Configuration procedure

Before the configuration, make sure that the device and the log host can reach each other.

1. Configure the device:

```
# Enable the information center.
```

```
<Sysname> system-view
```

```
[Sysname] info-center enable
```

```
# Specify the host 1.2.0.1/16 as the log host. Use channel loghost to output log information (optional, loghost by default), and use local4 as the logging facility.
```

```
[Sysname] info-center loghost 1.2.0.1 channel loghost facility local4
```

```
# Disable the output of log, trap, and debugging information of all modules on channel loghost.
```

```
[Sysname] info-center source default channel loghost debug state off log state off
trap state off
```

To avoid outputting unnecessary information, disable the output of log, trap, and debugging information on the specified channel (**loghost** in this example) before you configure an output rule.

Configure an output rule to output to the log host ARP and IP log information that has a severity level of at least **informational**. (The source modules that are allowed to output information depend on the switch model.)

```
[Sysname] info-center source arp channel loghost log level informational state on
[Sysname] info-center source ip channel loghost log level informational state on
```

2. Configure the log host:

The following configurations were performed on SunOS 4.0 which has similar configurations to the UNIX operating systems implemented by other vendors.

- a. Log in to the log host as a root user.
- b. Create a subdirectory named **Device** in directory **/var/log/**, and then create file **info.log** in the **Device** directory to save logs from **Device**.

```
# mkdir /var/log/Device
# touch /var/log/Device/info.log
```

- c. Edit the file **/etc/syslog.conf** and add the following contents.

```
# Device configuration messages
local4.info    /var/log/Device/info.log
```

In this configuration, **local4** is the name of the logging facility that the log host uses to receive logs. **info** is the information level. The UNIX system records the log information that has a severity of at least **informational** to the file **/var/log/Device/info.log**.

NOTE:

Be aware of the following issues while editing the file **/etc/syslog.conf**:

- Comments must be on a separate line and must begin with a pound sign (#).
- No redundant spaces are allowed after the file name.
- The logging facility name and the information level specified in the **/etc/syslog.conf** file must be identical to those configured on the device using the **info-center loghost** and **info-center source** commands. Otherwise the log information might not be output properly to the log host.

- d. Display the process ID of **syslogd**, kill the **syslogd** process and then restart **syslogd** using the **-r** option to make the modified configuration take effect.

```
# ps -ae | grep syslogd
147
# kill -HUP 147
# syslogd -r &
```

Now, the system can record log information into the log file.

Outputting log information to a Linux log host

Network requirements

Configure the device to send log information that has a severity level of at least informational to the Linux log host at 1.2.0.1/16.

Figure 19 Network diagram



Configuration procedure

Before the configuration, make sure that the device and the PC can reach each other.

1. Configure the device:

Enable the information center.

```
<Sysname> system-view
[Sysname] info-center enable
```

Specify the host 1.2.0.1/16 as the log host. Use the channel **loghost** to output log information (optional, **loghost** by default), and use **local5** as the logging facility.

```
[Sysname] info-center loghost 1.2.0.1 channel loghost facility local5
```

Disable the output of log, trap, and debugging information of all modules on channel **loghost**.

```
[Sysname] info-center source default channel loghost debug state off log state off
trap state off
```

To avoid outputting unnecessary information, disable the output of log, trap, and debugging information on the specified channel (**loghost** in this example) before you configure an output rule.

Configure an output rule to output to the log host the log information that has a severity level of at least **informational**.

```
[Sysname] info-center source default channel loghost log level informational state
on
```

2. Configure the log host:

a. Log in to the log host as a root user.

b. Create a subdirectory named **Device** in directory **/var/log/**, and create file **info.log** in the **Device** directory to save logs of **Device**.

```
# mkdir /var/log/Device
# touch /var/log/Device/info.log
```

c. Edit the file **/etc/syslog.conf** and add the following contents.

```
# Device configuration messages
local5.info    /var/log/Device/info.log
```

In this configuration, **local5** is the name of the logging facility that the log host uses to receive logs. The information level is **info**. The Linux system records the log information that has a severity level of at least **informational** to the file **/var/log/Device/info.log**.

NOTE:

Be aware of the following issues while editing the file **/etc/syslog.conf**:

- Comments must be on a separate line and must begin with a pound sign (#).
 - No redundant spaces are allowed after the file name.
 - The logging facility name and the information level specified in the **/etc/syslog.conf** file must be identical to those configured on the device using the **info-center loghost** and **info-center source** commands. Otherwise the log information may not be output properly to the log host.
-

- d. Display the process ID of **syslogd**, kill the **syslogd** process, and restart **syslogd** using the **-r** option to make the modified configuration take effect.

```
# ps -ae | grep syslogd
147
# kill -9 147
# syslogd -r &
```

Make sure that the **syslogd** process is started with the **-r** option on the Linux log host.

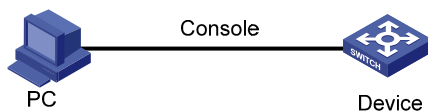
Now, the system can record log information into the log file.

Outputting log information to the console

Network requirements

Configure the device to send ARP and IP log information that has a severity level of at least Informational to the console.

Figure 20 Network diagram



Configuration procedure

```
# Enable the information center.
```

```
<Sysname> system-view
[Sysname] info-center enable
```

```
# Use channel console to output log information to the console. (This step is optional because it is the default setting).
```

```
[Sysname] info-center console channel console
```

```
# Disable the output of log, trap, and debugging information of all modules on channel console.
```

```
[Sysname] info-center source default channel console debug state off log state off trap
state off
```

To avoid outputting unnecessary information, disable the output of log, trap, and debugging information of all modules on the specified channel (**console** in this example), and then configure the output rule as needed.

```
# Configure an output rule to output to the console ARP and IP log information that has a severity level of at least informational. (The source modules that are allowed to output information depend on the switch model.)
```

```
[Sysname] info-center source arp channel console log level informational state on
[Sysname] info-center source ip channel console log level informational state on
[Sysname] quit
```

```
# Enable the display of log information on a terminal. (Optional, this function is enabled by default.)
```

```
<Sysname> terminal monitor
Info: Current terminal monitor is on.
<Sysname> terminal logging
Info: Current terminal logging is on.
```

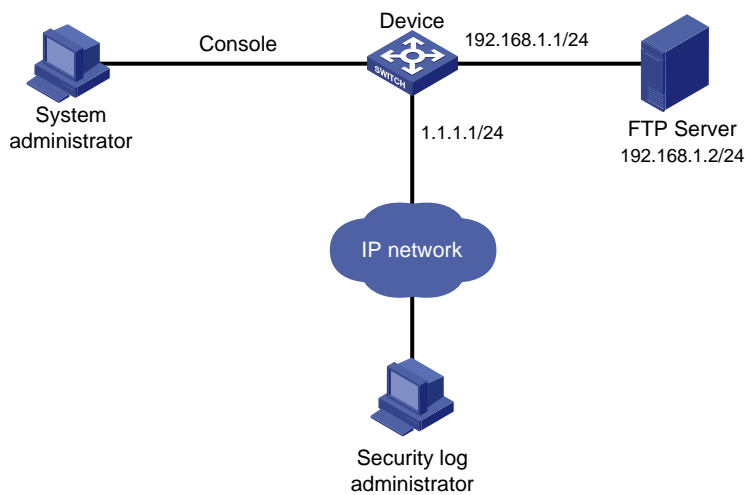
Now, if the ARP and IP modules generate log information, the information center automatically sends the log information to the console.

Saving security logs into the security log file

Network requirements

- Save security logs into the security log file **Flash:/securitylog/seclog.log** every one hour.
- Only the security log administrator can view the contents of the security log file. No other users cannot view, copy, or rename the security log file.

Figure 21 Network diagram



Configuration considerations

The configuration in this example includes two parts:

1. Log in to the device as the system administrator:
 - Enable saving the security logs into the security log file and set the saving interval to one hour.
 - Create a local user **seclog** with the password **123123123123**, and authorize this user as the security log administrator. That is, use the **authorization-attribute** command to set the user privilege level to 3 and specify the user role as security audit. In addition, specify the service types that the user can use by using **service-type**.
 - Set the authentication mode to **scheme** for the user logging in to the device, and make sure that only the local user that has passed the AAA local authentication can view and perform operations on the security log file.
2. Log in to the device as the security log administrator:
 - Set the directory for saving the security log file to **Flash:/securitylog/seclog.log**.
 - View the contents of the security log file to learn the security status of the device.

Configuration procedure

1. Configuration performed by the system administrator:

```
# Enable saving security logs into the security log file and set the saving interval to one hour.
<Sysname> system-view
```

```

[Sysname] info-center security-logfile enable
[Sysname] info-center security-logfile frequency 3600
# Create a local user seclog, and configure the password for the user as 123123123123.
[Sysname] local-user seclog
New local user added.
[Sysname-luser-seclog] password simple 123123123123
# Authorize the user to manage the security log file.
[Sysname-luser-seclog] authorization-attribute level 3 user-role security-audit
# Authorize the user to use SSH, Telnet, and terminal services.
[Sysname-luser-seclog] service-type ssh telnet terminal
[Sysname-luser-seclog] quit
# According to the network plan, the user logs in to the device through SSH or Telnet, so configure
the authentication mode of the VTY user interface as scheme.
[Sysname] display user-interface vty ?
    INTEGER<0-15> Specify one user terminal interface
The output shows that the device supports sixteen VTY user interfaces, which are numbered 0
through 15.
[Sysname] user-interface vty 0 15
[Sysname-ui-vty0-15] authentication-mode scheme
[Sysname-ui-vty0-15] quit

```

2. Configuration performed by the security log administrator:

```

# Log in to the device as user seclog.
C:/> telnet 1.1.1.1
*****
* Copyright (c) 2010-2012 Hewlett-Packard Development Company, L.P.          *
* Without the owner's prior written consent,                                *
* no decompiling or reverse-engineering shall be allowed.                  *
*****

```

Login authentication

```

Username:seclog
Password:
<Sysname>
# Display the summary of the security log file.
<Sysname> display security-logfile summary
  Security-log is enabled.
  Security-log file size quota: 1MB
  Security-log file directory: flash:/seclog
  Alarm-threshold: 80%
  Current usage: 0%
  Writing frequency: 1 hour 0 min 0 sec

```

The output shows that the directory for saving the security log file is **flash:/seclog**.

Change the directory where the security log file is saved to **Flash:/securitylog**.

```
<Sysname> mkdir securitylog
.
%Created dir flash:/securitylog.
<Sysname> info-center security-logfile switch-directory flash:/securitylog/
# Display the contents of the security log file buffer.
<Sysname> display security-logfile buffer
%@175 Nov  2 17:02:53:766 2011 Sysname SHELL/4/LOGOUT:
  Trap 1.3.6.1.4.1.25506.2.2.1.1.3.0.2: logout from Console
%@176 Nov  2 17:02:53:766 2011 Sysname SHELL/5/SHELL_LOGOUT:Console logged out from
aux0.
```

The content of other logs is not shown.

The preceding information indicates that there is still new content in the buffer that has not been saved into the security log file.

Manually save the contents of the security log file buffer into the security log file.

```
<Sysname> security-logfile save
Info: Save all the contents in the security log buffer into file
flash:/securitylog/seclog.log successfully.
```

Display the contents of the security log file.

```
<Sysname> more securitylog/seclog.log
%@157 Nov  2 16:12:01:750 2011 Sysname SHELL/4/LOGIN:
  Trap 1.3.6.1.4.1.25506.2.2.1.1.3.0.1: login from Console
%@158 Nov  2 16:12:01:750 2011 Sysname SHELL/5/SHELL_LOGIN:Console logged in from
aux0.
```

The content of other logs is not shown.

Configuring SNMP

This chapter provides an overview of the Simple Network Management Protocol (SNMP) and guides you through the configuration procedure.

Overview

SNMP is an Internet standard protocol widely used for a management station to access and operate the devices on a network, regardless of their vendors, physical characteristics and interconnect technologies.

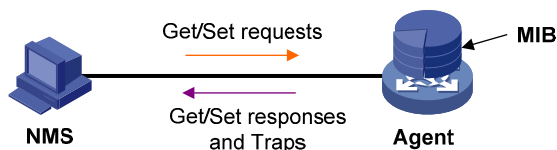
SNMP enables network administrators to read and set the variables on managed devices for state monitoring, troubleshooting, statistics collection, and other management purposes.

SNMP framework

The SNMP framework comprises the following elements:

- **SNMP manager**—Works on an NMS to monitor and manage the SNMP-capable devices in the network.
- **SNMP agent**—Works on a managed device to receive and handle requests from the NMS, and send traps to the NMS when some events, such as an interface state change, occur.
- **Management Information Base (MIB)**—Specifies the variables (for example, interface status and CPU usage) maintained by the SNMP agent for the SNMP manager to read and set.

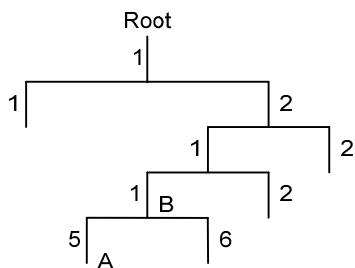
Figure 22 Relationship between an NMS, agent and MIB



MIB and view-based MIB access control

A MIB stores variables called "nodes" or "objects" in a tree hierarchy and identifies each node with a unique OID. An OID is a string of numbers that describes the path from the root node to a leaf node. For example, object B in Figure 23 is uniquely identified by the OID {1.2.1.1}.

Figure 23 MIB tree



A MIB view represents a set of MIB objects (or MIB object hierarchies) with certain access privilege and is identified by a view name. The MIB objects included in the MIB view are accessible while those excluded from the MIB view are inaccessible.

A MIB view can have multiple view records each identified by a *view-name oid-tree* pair.

You control access to the MIB by assigning MIB views to SNMP groups or communities.

SNMP operations

SNMP provides the following basic operations:

- **Get**—The NMS retrieves SNMP object nodes in an agent MIB.
- **Set**—The NMS modifies the value of an object node in an agent MIB.
- **Notifications**—Includes traps and informs. SNMP agent sends traps or informs to report events to the NMS. The difference between these two types of notification is that informs require acknowledgement but traps do not. The device supports only traps.

SNMP protocol versions

HP supports SNMPv1, SNMPv2c, and SNMPv3. An NMS and an SNMP agent must use the same SNMP version to communicate with each other.

- **SNMPv1**—Uses community names for authentication. To access an SNMP agent, an NMS must use the same community name as set on the SNMP agent. If the community name used by the NMS is different from that set on the agent, the NMS cannot establish an SNMP session to access the agent or receive traps from the agent.
- **SNMPv2c**—Uses community names for authentication. SNMPv2c is compatible with SNMPv1, but supports more operation modes, data types, and error codes.
- **SNMPv3**—Uses a user-based security model (USM) to secure SNMP communication. You can configure authentication and privacy mechanisms to authenticate and encrypt SNMP packets for integrity, authenticity, and confidentiality.

SNMP configuration task list

| Task | Remarks |
|---|----------|
| Configuring SNMP basic parameters | Required |
| Switching the NM-specific interface index | Optional |
| Configuring SNMP logging | Optional |
| Configuring SNMP traps | Optional |

Configuring SNMP basic parameters

SNMPv3 differs from SNMPv1 and SNMPv2c in many ways. Their configuration procedures are described in separate sections.

Configuring SNMPv3 basic parameters

| Step | Command | Remarks |
|---|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enable the SNMP agent. | snmp-agent | Optional.
By default, the SNMP agent is disabled.

You can also enable the SNMP agent by using any command that begins with snmp-agent except the snmp-agent calculate-password and snmp-agent ifmib long-ifindex enable commands. |
| 3. Configure system information for the SNMP agent. | snmp-agent sys-info { contact <i>sys-contact</i> location <i>sys-location</i> version { all { v1 v2c v3 }* } } | Optional.
By default, the contact information is Hewlett-Packard Development Company, L.P. , the location information is null, and the protocol version is SNMPv3 . |
| 4. Configure the local engine ID. | snmp-agent local-engineid <i>engineid</i> | Optional.
The default local engine ID is the company ID plus the device ID.
After you change the local engine ID, the existing SNMPv3 users become invalid, and you must re-create the SNMPv3 users. |
| 5. Create or update a MIB view. | snmp-agent mib-view { excluded included } <i>view-name oid-tree</i> [mask <i>mask-value</i>] | Optional.
By default, the MIB view ViewDefault is predefined and its OID is 1.

Each <i>view-name oid-tree</i> pair represents a view record. If you specify the same record with different MIB subtree masks multiple times, the last configuration takes effect. Except the four subtrees in the default MIB view, you can create up to 16 unique MIB view records. |
| 6. Configure an SNMPv3 group. | snmp-agent group v3 <i>group-name</i> [authentication privacy] [read-view <i>read-view</i>] [write-view <i>write-view</i>] [notify-view <i>notify-view</i>] [acl <i>acl-number</i> acl ipv6 <i>ipv6-acl-number</i>] * | By default, no SNMP group exists. |
| 7. Convert a plaintext key to a ciphertext (encrypted) key. | snmp-agent calculate-password <i>plain-password</i> mode { 3desmd5 3dessha md5 sha } { local-engineid specified-engineid <i>engineid</i> } | Optional. |

| Step | Command | Remarks |
|--|---|---|
| 8. Add a user to the SNMPv3 group. | snmp-agent usm-user v3
<i>user-name group-name</i> [[cipher]
authentication-mode { md5 sha }
<i>auth-password</i> [privacy-mode
{ 3des aes128 des56 }
<i>priv-password</i>]] [acl <i>acl-number</i>
acl ipv6 <i>ipv6-acl-number</i>] * | N/A |
| 9. Configure the maximum SNMP packet size (in bytes) that the SNMP agent can handle. | snmp-agent packet max-size
<i>byte-count</i> | Optional.
By default, the SNMP agent can receive and send SNMP packets up to 1500 bytes. |
| 10. Configure the DSCP value for SNMP responses. | snmp-agent packet response dscp
<i>dscp-value</i> | Optional.
By default, the DSCP value for SNMP responses is 0. |

Configuring SNMPv1 or SNMPv2c basic parameters

| Step | Command | Remarks |
|---|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enable the SNMP agent. | snmp-agent | Optional.
By default, the SNMP agent is disabled.
You can also enable the SNMP agent service by using any command that begins with snmp-agent except the snmp-agent calculate-password and snmp-agent ifmib long-ifindex enable commands. |
| 3. Configure system information for the SNMP agent. | snmp-agent sys-info { contact <i>sys-contact</i>
location <i>sys-location</i> version { all { v1 v2c v3 }* } } | By default, the contact information is Hewlett-Packard Development Company, L.P. , the location information is null, and the protocol version is SNMPv3 . |
| 4. Configure the local engine ID. | snmp-agent local-engineid <i>engineid</i> | Optional.
The default local engine ID is the company ID plus the device ID. |

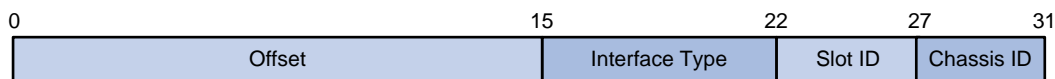
| Step | Command | Remarks |
|--|---|--|
| 5. Create or update a MIB view. | snmp-agent mib-view { excluded included }
<i>view-name oid-tree</i> [mask mask-value] | Optional.
By default, the MIB view ViewDefault is predefined and its OID is 1.
Each <i>view-name oid-tree</i> pair represents a view record. If you specify the same record with different MIB subtree masks multiple times, the last configuration takes effect. Except the four subtrees in the default MIB view, you can create up to 16 unique MIB view records. |
| 6. Configure the SNMP access right. | <ul style="list-style-type: none"> • (Approach 1) Create an SNMP community:
snmp-agent community { read write }
<i>community-name</i> [mib-view view-name]
[acl acl-number acl ipv6 ipv6-acl-number] * • (Approach 2) Create an SNMP group, and add a user to the SNMP group: <ul style="list-style-type: none"> a. snmp-agent group { v1 v2c }
<i>group-name</i> [read-view read-view]
[write-view write-view] [notify-view notify-view] [acl acl-number acl ipv6 ipv6-acl-number] * b. snmp-agent usm-user { v1 v2c }
<i>user-name group-name</i> [acl acl-number acl ipv6 ipv6-acl-number] * | Use either approach.
By default, no SNMP group exists.
In approach 2, the username is equivalent to the community name in approach 1, and must be the same as the community name configured on the NMS. |
| 7. Configure the maximum size (in bytes) of SNMP packets for the SNMP agent. | snmp-agent packet max-size <i>byte-count</i> | Optional.
By default, the SNMP agent can receive and send the SNMP packets up to 1500 bytes. |
| 8. Configure the DSCP value for SNMP responses. | snmp-agent packet response dscp <i>dscp-value</i> | Optional.
By default, the DSCP value for SNMP responses is 0. |

Switching the NM-specific interface index format

A network management (NM)-specific ifindex identifies an interface and is provided by the SNMP managed device to the NMS. A network management-specific ifindex takes one of the following two formats:

- **16-bit NM-specific ifindex**—The system dynamically assigns 16-bit NM-specific ifindex values to uniquely identify its interfaces. The 16-bit NM-specific ifindex value starts from 1 and increments by 1.
- **32-bit NM-specific ifindex**—A 32-bit NM-specific ifindex value comprises an Offset, Interface Type, Slot ID, and Chassis ID, as shown in [Figure 24](#).

Figure 24 32-bit NM-specific ifindex



- **Offset**—This field is 16 bits long and distinguishes different interfaces of the same type on the same interface card.
- **Interface type**—This field is 7 bits long and contains the enumerated value specific to the interface type. It supports up to 128 different interface types and supports more than 80 interface types at present.
- **Slot ID**—This field is 5 bits long and contains the number of the physical slot that holds the interface.
- **Chassis ID**—This field is 4 bits long. For a distributed device in IRF mode, this field indicates the member ID of the device that provides the interface. For other types of devices, this field has no meanings and the value is 0.

Configuration guidelines

- Use the 32-bit NM-specific ifindex format if the NMS requires the format to get information such as the slot that contains a specific interface. If the network protocol operating on the NMS does not support 32-bit NM-specific ifindex values, make sure NM-specific ifindex values on the device are 16-bit. By default, the device adopts the 16-bit NM-specific ifindex format.
- An NM-specific ifindex format change invalidates the NM-specific ifindex dependent settings, and these settings cannot become valid until you switch the format back. To use these settings in the new format, you must re-configure them. For example, if an RMON alarm group or private alarm group has alarm variables in the format *OID/variable-name.NM-specific-ifindex*, you must reconfigure these variables after an NM-specific ifindex format change.

Configuration procedure

To switch the NM-specific ifindex format:

| Step | Command | Remarks |
|---|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Switch the format of an NM-specific ifindex from 16-bit to 32-bit. | snmp-agent ifmib long-ifindex enable | Optional.
By default, an NM-specific ifindex is in 16-bit format. |

| Step | Command | Remarks |
|------|--|--|
| 3. | Switch the format of an NM-specific ifindex from 32-bit to 16-bit.
undo snmp-agent ifmib long-ifindex enable | Optional.
By default, an NM-specific ifindex is in 16-bit format. |

Configuring SNMP logging

ⓘ IMPORTANT:

Disable SNMP logging in normal cases to prevent a large amount of SNMP logs from decreasing device performance.

The SNMP logging function logs Get requests, Set requests, and Set responses, but does not log Get responses.

- **Get operation**—The agent logs the IP address of the NMS, name of the accessed node, and node OID.
- **Set operation**—The agent logs the NMS' IP address, name of accessed node, node OID, and error code and index for the Set operation.

The SNMP module sends these logs to the information center as informational messages. You can configure the information center to output these messages to certain destinations, for example, the console and the log buffer. The total output size for the node field (MIB node name) and the value field (value of the MIB node) in each log entry is 1024 bytes. If this limit is exceeded, the information center truncates the data in the fields. For more information about the information center, see "[Configuring the information center](#)."

To configure SNMP logging:

| Step | Command | Remarks |
|------|---|---------------------------------------|
| 1. | Enter system view.
system-view | N/A |
| 2. | Enable SNMP logging.
snmp-agent log { all get-operation set-operation } | By default, SNMP logging is disabled. |

Configuring SNMP traps

The SNMP agent sends traps to inform the NMS of important events, such as a reboot.

Traps fall into generic traps and vendor-specific traps. Generic traps include **authentication**, **coldstart**, **linkdown**, **linkup** and **warmstart**. All other traps are vendor-defined.

SNMP traps generated by a module are sent to the information center. You can configure the information center to enable or disable outputting the traps from a module by severity and set output destinations. For more information about the information center, see "[Configuring the information center](#)."

Enabling SNMP traps

Enable SNMP traps only if necessary. SNMP traps are memory-intensive and may affect device performance.

To generate linkUp or linkDown traps when the link state of an interface changes, you must enable the linkUp or linkDown trap function globally by using the **snmp-agent trap enable [standard [linkdown | linkup] *]** command and on the interface by using the **enable snmp trap updown** command.

After you enable a trap function for a module, whether the module generates traps also depends on the configuration of the module. For more information, see the configuration guide for each module.

To enable traps:

| Step | Command | Remarks |
|--|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enable traps globally. | snmp-agent trap enable [arp rate-limit bgp configuration default-route flash ospf [process-id] [ifauthfail ifcfgerror ifrxbadpkt ifstatechange iftxretransmit lsdapproachoverflow lsdoverflow maxagelsa nbrstatechange originatelsa vifcfgerror virifauthfail virifrxbadpkt virifstatechange viriftxretransmit virnbrstatechange] * pim [candidatebsrwinelection electedbsrlostelection interfaceelection invalidjoinprune invalidregister neighborloss rpmappingchange] * standard [authentication coldstart linkdown linkup warmstart] * system vrrp [authfailure newmaster]] | By default, all traps are enabled.
The HP 5500 SI switches do not support the bgp , ospf , pim , or vrrp keyword. |
| 3. Enter Layer 2 Ethernet interface view, Layer 3 Ethernet interface view, or VLAN interface view. | interface interface-type interface-number | Only the HP 5500 EI switches support Layer 3 Ethernet port configuration. |
| 4. Enable link state traps. | enable snmp trap updown | By default, the link state traps are enabled. |

Configuring the SNMP agent to send traps to a host

The SNMP module buffers the traps received from a module in a trap queue. You can set the size of the queue, the duration that the queue holds a trap, and trap target (destination) hosts, typically the NMS.

To successfully send traps, you must also perform the following tasks:

- Complete the basic SNMP settings and verify that they are the same as on the NMS. If SNMPv1 or SNMPv2c is used, you must configure a community name. If SNMPv3 is used, you must configure an SNMPv3 user and MIB view.
- Make sure the device and the NMS can reach each other.

To configure the SNMP agent to send traps to a host:

| Step | Command | Remarks |
|-----------------------|--------------------|---------|
| 1. Enter system view. | system-view | N/A |

| Step | Command | Remarks |
|---|--|---|
| 2. Configure a target host. | snmp-agent target-host trap address udp-domain { <i>ip-address</i> ipv6 <i>ipv6-address</i> } [udp-port <i>port-number</i>] [dscp <i>dscp-value</i>] [vpn-instance <i>vpn-instance-name</i>]
params securityname <i>security-string</i> [v1 v2c v3] [authentication privacy]] | If the trap destination is a host, the <i>ip-address</i> argument must be the IP address of the host.
Only the HP 5500 EI switches support the vpn-instance keyword. |
| 3. Configure the source address for traps. | snmp-agent trap source interface-type interface-number | Optional.
By default, SNMP chooses the IP address of an interface to be the source IP address of traps. |
| 4. Extend the standard linkUp/linkDown traps. | snmp-agent trap if-mib link extended | Optional.
By default, standard linkUp/linkDown traps are used.
Extended linkUp/linkDown traps add interface description and interface type to standard linkUp/linkDown traps. If the NMS does not support extended SNMP messages, use standard linkUp/linkDown traps. |
| 5. Configure the trap queue size. | snmp-agent trap queue-size <i>size</i> | Optional.
The default trap queue size is 100.
When the trap queue is full, the oldest traps are automatically deleted for new traps. |
| 6. Configure the trap holding time. | snmp-agent trap life <i>seconds</i> | Optional.
The default setting is 120 seconds.
A trap is deleted when its holding time expires. |

Displaying and maintaining SNMP

| Task | Command | Remarks |
|--|--|-----------------------|
| Display SNMP agent system information, including the contact, physical location, and SNMP version. | display snmp-agent sys-info [contact location version]* [[{ begin exclude include } <i>regular-expression</i>]] | Available in any view |
| Display SNMP agent statistics. | display snmp-agent statistics [[{ begin exclude include } <i>regular-expression</i>]] | Available in any view |
| Display the local engine ID. | display snmp-agent local-engineid [[{ begin exclude include } <i>regular-expression</i>]] | Available in any view |
| Display SNMP group information. | display snmp-agent group [<i>group-name</i>] [[{ begin exclude include } <i>regular-expression</i>]] | Available in any view |

| Task | Command | Remarks |
|--|---|-----------------------|
| Display basic information about the trap queue. | display snmp-agent trap queue [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display the modules that can send traps and their trap status (enable or disable). | display snmp-agent trap-list [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display SNMPv3 user information. | display snmp-agent usm-user [engineid <i>engineid</i> username <i>user-name</i> group <i>group-name</i>] * [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display SNMPv1 or SNMPv2c community information. | display snmp-agent community [read write] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display MIB view information. | display snmp-agent mib-view [exclude include viewname <i>view-name</i>] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |

SNMP configuration examples

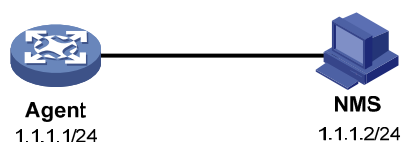
This section gives examples of how to configure SNMPv1 or SNMPv2c, SNMPv3, and SNMP logging.

SNMPv1/SNMPv2c configuration example

Network requirements

As shown in [Figure 25](#), the NMS (1.1.1.2/24) uses SNMPv1 or SNMPv2c to manage the SNMP agent (1.1.1.1/24), and the agent automatically sends traps to report events to the NMS.

Figure 25 Network diagram



Configuration procedure

1. Configure the SNMP agent:

Configure the IP address of the agent, and make sure the agent and the NMS can reach each other. (Details not shown.)

Specify SNMPv1 and SNMPv2c, and create a read-only community **public** and a read and write community **private**.

```

<Agent> system-view
[Agent] snmp-agent sys-info version v1 v2c
[Agent] snmp-agent community read public
[Agent] snmp-agent community write private
  
```

Configure contact and physical location information for the agent.

```

[Agent] snmp-agent sys-info contact Mr.Wang-Tel:3306
[Agent] snmp-agent sys-info location telephone-closet,3rd-floor
  
```

```
# Enable SNMP traps, set the NMS at 1.1.1.2 as an SNMP trap destination, and use public as the
community name. (To make sure the NMS can receive traps, specify the same SNMP version in the
snmp-agent target-host command as is configured on the NMS.)
```

```
[Agent] snmp-agent trap enable
```

```
[Agent] snmp-agent target-host trap address udp-domain 1.1.1.2 params securityname
public v1
```

```
[Agent] quit
```

2. Configure the SNMP NMS:

```
# Configure the SNMP version for the NMS as v1 or v2c, create a read-only community and name
it public, and create a read and write community and name it private. For information about
configuring the NMS, see the NMS manual.
```

NOTE:

The SNMP settings on the agent and the NMS must match.

3. Verify the configuration:

```
# Try to get the count of sent traps from the agent. The attempt succeeds.
```

```
Send request to 1.1.1.1/161 ...
```

```
Protocol version: SNMPv1
```

```
Operation: Get
```

```
Request binding:
```

```
1: 1.3.6.1.2.1.11.29.0
```

```
Response binding:
```

```
1: Oid=snmpOutTraps.0 Syntax=CNTR32 Value=18
```

```
Get finished
```

```
# Use a wrong community name to get the value of a MIB node from the agent. You can see an
authentication failure trap on the NMS.
```

```
1.1.1.1/2934 V1 Trap = authenticationFailure
```

```
SNMP Version = V1
```

```
Community = public
```

```
Command = Trap
```

```
Enterprise = 1.3.6.1.4.1.43.1.16.4.3.50
```

```
GenericID = 4
```

```
SpecificID = 0
```

```
Time Stamp = 8:35:25.68
```

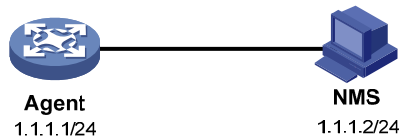
SNMPv3 configuration example

Network requirements

As shown in [Figure 26](#), the NMS (1.1.1.2/24) uses SNMPv3 to monitor and manage the interface status of the agent (1.1.1.1/24), and the agent automatically sends traps to report events to the NMS.

The NMS and the agent perform authentication when they set up an SNMP session. The authentication algorithm is MD5 and the authentication key is **authkey**. The NMS and the agent also encrypt the SNMP packets between them by using the DES algorithm and the privacy key **prikey**.

Figure 26 Network diagram



Configuration procedure

1. Configure the agent:

Configure the IP address of the agent and make sure the agent and the NMS can reach each other. (Details not shown.)

Assign the NMS read and write access to the objects under the snmp node (OID 1.3.6.1.2.1.11), and deny its access to any other MIB object.

```
<Agent> system-view
[Agent] undo snmp-agent mib-view ViewDefault
[Agent] snmp-agent mib-view included test snmp
[Agent] snmp-agent group v3 managev3group read-view test write-view test
```

Set the username to **managev3user**, authentication algorithm to **MD5**, authentication key to **authkey**, encryption algorithm to **DES56**, and privacy key to **prikey**.

```
[Agent] snmp-agent usm-user v3 managev3user managev3group authentication-mode md5
authkey privacy-mode des56 prikey
```

Configure contact person and physical location information for the agent.

```
[Agent] snmp-agent sys-info contact Mr.Wang-Tel:3306
[Agent] snmp-agent sys-info location telephone-closet,3rd-floor
```

Enable traps, specify the NMS at 1.1.1.2 as a trap destination, and set the username to **managev3user** for the traps.

```
[Agent] snmp-agent trap enable
[Agent] snmp-agent target-host trap address udp-domain 1.1.1.2 params securityname
managev3user v3 privacy
```

2. Configure the SNMP NMS:

- Specify the SNMP version for the NMS as v3.
- Create two SNMP users: **managev3user** and **public**.
- Enable both authentication and privacy functions.
- Use MD5 for authentication and DES for encryption.
- Set the authentication key to **authkey** and the privacy key to **prikey**.
- Set the timeout time and maximum number of retries.

For information about configuring the NMS, see the NMS manual.

NOTE:

The SNMP settings on the agent and the NMS must match.

3. Verify the configuration:

Try to get the count of sent traps from the agent. The get attempt succeeds.

```
Send request to 1.1.1.1/161 ...
Protocol version: SNMPv3
Operation: Get
```

```

Request binding:
1: 1.3.6.1.2.1.11.29.0
Response binding:
1: Oid=snmpOutTraps.0 Syntax=CNTR32 Value=18
Get finished
# Try to get the device name from the agent. The get attempt fails because the NMS has no access
right to the node.
Send request to 1.1.1.1/161 ...
Protocol version: SNMPv3
Operation: Get
Request binding:
1: 1.3.6.1.2.1.1.5.0
Response binding:
1: Oid=sysName.0 Syntax=noSuchObject Value=NULL
Get finished
# Execute the shutdown or undo shutdown command on an idle interface on the agent. You can
see the interface state change traps on the NMS:
1.1.1.1/3374 V3 Trap = linkdown
SNMP Version = V3
Community = managev3user
Command = Trap
1.1.1.1/3374 V3 Trap = linkup
SNMP Version = V3
Community = managev3user
Command = Trap

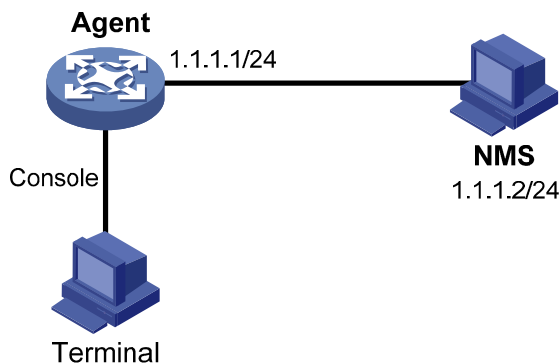
```

SNMP logging configuration example

Network requirements

Configure the SNMP agent (1.1.1.1/24) in [Figure 27](#) to log the SNMP operations performed by the NMS.

Figure 27 Network diagram



Configuration procedure

This example assumes you have configured all required SNMP settings for the NMS and the agent (see "[SNMPv1/SNMPv2c configuration example](#)" or "[SNMPv3 configuration example](#)").

Enable displaying log messages on the configuration terminal. (This function is enabled by default. Skip this step if you are using the default.)

```
<Agent> terminal monitor
<Agent> terminal logging
```

Enable the information center to output system information with severity level equal to or higher than informational to the console port.

```
<Agent> system-view
[Agent] info-center source snmp channel console log level informational
```

Enable logging GET and SET operations.

```
[Agent] snmp-agent log all
```

Verify the configuration:

Use the NMS to get a MIB variable from the agent. The following is a sample log message displayed on the configuration terminal:

```
%Jan 1 02:49:40:566 2011 Sysname SNMP/6/GET:
seqNO = <10> srcIP = <1.1.1.2> op = <get> node = <sysName(1.3.6.1.2.1.1.5.0)> value=<>
```

Use the NMS to set a MIB variable on the agent. The following is a sample log message displayed on the configuration terminal:

```
%Jan 1 02:59:42:576 2011 Sysname SNMP/6/SET:
seqNO = <11> srcIP = <1.1.1.2> op = <set> errorIndex = <0> errorStatus =<noError> node
= <sysName(1.3.6.1.2.1.1.5.0)> value = <Agent>
```

Table 6 SNMP log message field description

| Field | Description |
|-------------------------|--|
| Jan 1 02:49:40:566 2011 | Time when the SNMP log was generated. |
| seqNO | Serial number automatically assigned to the SNMP log, starting from 0. |
| srcIP | IP address of the NMS. |
| op | SNMP operation type (GET or SET). |
| node | MIB node name and OID of the node instance. |
| errorIndex | Error index, with 0 meaning no error. |
| errorStatus | Error status, with noError meaning no error. |
| value | Value set by the SET operation. This field is null for a GET operation.
If the value is a character string that has invisible characters or characters beyond the ASCII range 0 to 127, the string is displayed in hexadecimal format, for example, value = <81-43>[hex]. |

The information center can output system event messages to several destinations, including the terminal and the log buffer. In this example, SNMP log messages are output to the terminal. To configure other message destinations, see "[Configuring the information center.](#)"

Configuring RMON

This chapter describes how to configure RMON.

Overview

Remote Monitoring (RMON) is an enhancement to SNMP for remote device management and traffic monitoring. An RMON monitor, typically the RMON agent embedded in a network device, periodically or continuously collects traffic statistics for the network attached to a port, and when a statistic crosses a threshold, logs the crossing event and sends a trap to the management station.

RMON uses SNMP traps to notify NMSs of exceptional conditions. RMON SNMP traps report various events, including traffic events such as broadcast traffic threshold exceeded. In contrast, SNMP standard traps report device operating status changes such as link up, link down, and module failure.

RMON enables proactive monitoring and management of remote network devices and subnets. The managed device can automatically send a trap when a statistic crosses an alarm threshold, and the NMS does not need to constantly poll MIB variables and compare the results. As a result, network traffic is reduced.

Working mechanism

RMON monitors typically take one of the following forms:

- **Dedicated RMON probes.** NMSs can obtain management information from RMON probes directly and control network resources. In this approach, NMSs can obtain all RMON MIB information.
- **RMON agents embedded in network devices.** NMSs exchange data with RMON agents by using basic SNMP operations to gather network management information. Because this approach is resource intensive, most RMON agent implementations provide only four groups of MIB information: alarm, event, history, and statistics.

HP devices provide the embedded RMON agent function. You can configure your device to collect and report traffic statistics, error statistics, and performance statistics.

RMON groups

Among the RFC 2819 defined RMON groups, HP implements the statistics group, history group, event group, and alarm group supported by the public MIB. HP also implements a private alarm group, which enhances the standard alarm group.

Ethernet statistics group

The statistics group defines that the system collects traffic statistics on interfaces (only Ethernet interfaces are supported) and saves the statistics in the Ethernet statistics table (ethernetStatsTable). The interface traffic statistics include network collisions, CRC alignment errors, undersize/oversize packets, broadcasts, multicasts, bytes received, and packets received.

After you create a statistics entry for an interface, the statistics group starts to collect traffic statistics on the interface. The statistics in the Ethernet statistics table are cumulative sums.

History group

The history group defines that the system periodically collects traffic statistics on interfaces and saves the statistics in the history record table (ethernetHistoryTable). The statistics include bandwidth utilization, number of error packets, and total number of packets.

The history statistics table record traffic statistics collected for each sampling interval. The sampling interval is user-configurable.

Event group

The event group defines event indexes and controls the generation and notifications of the events triggered by the alarms defined in the alarm group and the private alarm group. The events can be handled in one of the following ways:

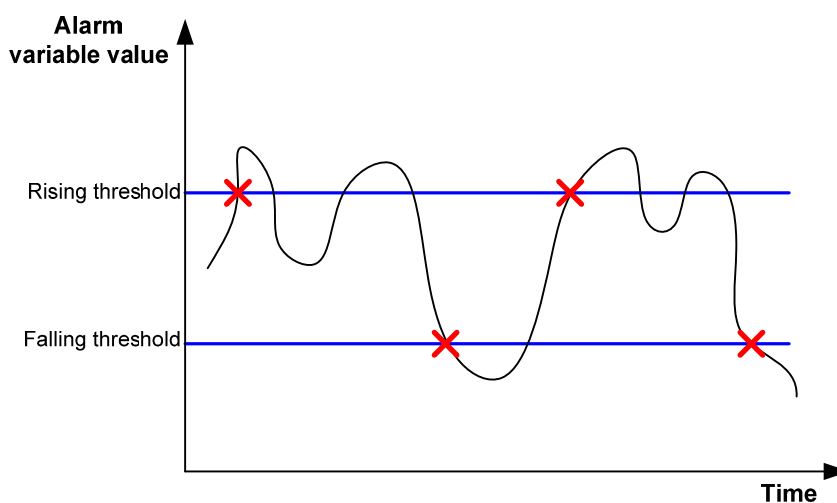
- **Log**—Logs event information (including event name and description) in the event log table of the RMON MIB, so the management device can get the logs through the SNMP Get operation.
- **Trap**—Sends a trap to notify an NMS of the event.
- **Log-Trap**—Logs event information in the event log table and sends a trap to the NMS.
- **None**—No action.

Alarm group

The RMON alarm group monitors alarm variables, such as the count of incoming packets (etherStatsPkts) on an interface. After you define an alarm entry, the system gets the value of the monitored alarm variable at the specified interval. If the value of the monitored variable is greater than or equal to the rising threshold, a rising event is triggered. If the value of the monitored variable is smaller than or equal to the falling threshold, a falling event is triggered. The event is then handled as defined in the event group.

If an alarm entry crosses a threshold multiple times in succession, the RMON agent generates an alarm event only for the first crossing. For example, if the value of a sampled alarm variable crosses the rising threshold multiple times before it crosses the falling threshold, only the first crossing triggers a rising alarm event, as shown in [Figure 28](#).

Figure 28 Rising and falling alarm events



Private alarm group

The private alarm group calculates the values of alarm variables and compares the results with the defined threshold for a more comprehensive alarming function.

The system handles the private alarm entry (as defined by the user) in the following ways:

- Periodically samples the prialarm variables defined in the prialarm formula.
- Calculates the sampled values based on the prialarm formula.
- Compares the result with the defined threshold and generates an appropriate event if the threshold value is reached.

If a private alarm entry crosses a threshold multiple times in succession, the RMON agent generates an alarm event only for the first crossing. For example, if the value of a sampled alarm variable crosses the rising threshold multiple times before it crosses the falling threshold, only the first crossing triggers a rising alarm event. If the count result of the private alarm group overpasses the same threshold multiple times, only the first one can cause an alarm event. In other words, the rising alarm and falling alarm are alternate.

Configuring the RMON statistics function

The RMON statistics function can be implemented by either the Ethernet statistics group or the history group, but the objects of the statistics are different, as follows:

- A statistics object of the Ethernet statistics group is a variable defined in the Ethernet statistics table, and the recorded content is a cumulative sum of the variable from the time the statistics entry is created to the current time. For more information, see "[Configuring the RMON Ethernet statistics function](#)."
- A statistics object of the history group is the variable defined in the history record table, and the recorded content is a cumulative sum of the variable in each period. For more information, see "[Configuring the RMON history statistics function](#)."

Configuring the RMON Ethernet statistics function

| Step | Command | Remarks |
|--|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter Layer 2 Ethernet interface view or Layer 3 Ethernet interface view. | interface <i>interface-type</i>
<i>interface-number</i> | You can configure an Ethernet port as a Layer 3 Ethernet port only on the HP 5500 EI switch. |
| 3. Create an entry in the RMON statistics table. | rmon statistics <i>entry-number</i> [owner <i>text</i>] | N/A |

You can create one statistics entry per interface and up to 100 statistics entries on the device. When the number of statistics entries exceeds 100, you cannot add new entries.

Configuring the RMON history statistics function

Follow these guidelines when you configure the RMON history statistics function:

- The *entry-number* for an RMON history control entry must be globally unique. If an entry number has been used on one interface, it cannot be used on another.

- You can configure multiple history control entries for one interface, but must make sure their entry numbers and sampling intervals are different.
- On the HP 5500 SI switch, up to 100 history entries can be created. On the HP 5500 EI switch, up to 1000 history entries can be created.
- You can successfully create a history control entry, even if the specified bucket size exceeds the history table size supported by the device. However, the effective bucket size will be the actual value supported by the device.

To configure the RMON history statistics function:

| Step | Command | Remarks |
|--|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter Layer 2 Ethernet interface view or Layer 3 Ethernet interface view. | interface <i>interface-type</i>
<i>interface-number</i> | You can configure an Ethernet port as a Layer 3 Ethernet port only on the HP 5500 EI switch. |
| 3. Create an entry in the RMON history control table. | rmon history <i>entry-number</i> buckets
<i>number</i> interval <i>sampling-interval</i>
[owner <i>text</i>] | N/A |

Configuring the RMON alarm function

Follow these guidelines when you configure the RMON alarm function:

- To send traps to the NMS when an alarm is triggered, configure the SNMP agent as described in "Configuring SNMP" before configuring the RMON alarm function.
- If the alarm variable is a MIB variable defined in the history group or the Ethernet statistics group, make sure the RMON Ethernet statistics function or the RMON history statistics function is configured on the monitored Ethernet interface. Otherwise, even if you can create the alarm entry, no alarm event can be triggered.
- You cannot create a new event, alarm, or private alarm entry that has the same set of parameters as an existing entry. For parameters to be compared for duplication, see [Table 7](#).
- After the maximum number of entries is reached, no new entry can be created. For the table entry limits, see [Table 7](#).

To configure the RMON alarm function:

| Step | Command | Remarks |
|--|---|---------|
| 1. Enter system view. | system-view | N/A |
| 2. Create an event entry in the event table. | rmon event <i>entry-number</i> [description <i>string</i>] { log
log-trap <i>log-trapcommunity</i> none trap
<i>trap-community</i> } [owner <i>text</i>] | N/A |

| Step | Command | Remarks |
|---|---|---------------------------|
| 3. Create an entry in the alarm table or private alarm table. | <ul style="list-style-type: none"> Create an entry in the alarm table:
 rmon alarm <i>entry-number alarm-variable sampling-interval</i> { absolute delta }
 rising-threshold <i>threshold-value1 event-entry1</i>
 falling-threshold <i>threshold-value2 event-entry2</i>
 [owner text] Create an entry in the private alarm table:
 rmon prialarm <i>entry-number prialarm-formula prialarm-des sampling-interval</i> { absolute changeratio delta } rising-threshold <i>threshold-value1 event-entry1</i> falling-threshold <i>threshold-value2 event-entry2</i> entrytype { forever cycle <i>cycle-period</i> } [owner text] | Use at least one command. |

Table 7 RMON configuration restrictions

| Entry | Parameters to be compared | Maximum number of entries |
|----------|--|---------------------------|
| Event | Event description (description string), event type (log, trap, logtrap or none) and community name (<i>trap-community</i> or <i>log-trapcommunity</i>) | 60 |
| Alarm | Alarm variable (<i>alarm-variable</i>), sampling interval (<i>sampling-interval</i>), sampling type (absolute or delta), rising threshold (<i>threshold-value1</i>) and falling threshold (<i>threshold-value2</i>) | 60 |
| Prialarm | Alarm variable formula (<i>alarm-variable</i>), sampling interval (<i>sampling-interval</i>), sampling type (absolute, changeratio or delta), rising threshold (<i>threshold-value1</i>) and falling threshold (<i>threshold-value2</i>) | 50 |

Displaying and maintaining RMON

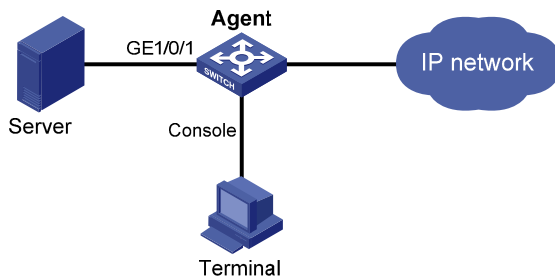
| Task | Command | Remarks |
|--|--|-----------------------|
| Display RMON statistics. | display rmon statistics [<i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display the RMON history control entry and history sampling information. | display rmon history [<i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display RMON alarm configuration. | display rmon alarm [<i>entry-number</i>] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display RMON private alarm configuration. | display rmon prialarm [<i>entry-number</i>] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display RMON events configuration. | display rmon event [<i>entry-number</i>] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display log information for event entries. | display rmon eventlog [<i>entry-number</i>] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |

Ethernet statistics group configuration example

Network requirements

Configure the RMON statistics group on the RMON agent in [Figure 29](#) to gather cumulative traffic statistics for GigabitEthernet 1/0/1.

Figure 29 Network diagram



Configuration procedure

Configure the RMON statistics group on the RMON agent to gather statistics for GigabitEthernet 1/0/1.

```
<Sysname> system-view
[Sysname] interface gigabitethernet 1/0/1
[Sysname-GigabitEthernet1/0/1] rmon statistics 1 owner user1
```

Display statistics collected by the RMON agent for GigabitEthernet 1/0/1.

```
<Sysname> display rmon statistics gigabitethernet 1/0/1
EtherStatsEntry 1 owned by user1-rmon is VALID.
  Interface : GigabitEthernet1/0/1<ifIndex.3>
  etherStatsOctets      : 21657      , etherStatsPkts      : 307
  etherStatsBroadcastPkts : 56      , etherStatsMulticastPkts : 34
  etherStatsUndersizePkts : 0      , etherStatsOversizePkts : 0
  etherStatsFragments   : 0      , etherStatsJabbers     : 0
  etherStatsCRCAlignErrors : 0      , etherStatsCollisions  : 0
  etherStatsDropEvents (insufficient resources): 0
  Packets received according to length:
  64      : 235      , 65-127 : 67      , 128-255 : 4
  256-511: 1      , 512-1023: 0      , 1024-1518: 0
```

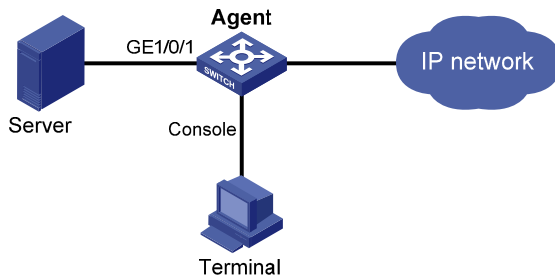
On the configuration terminal, get the traffic statistics through SNMP. (Details not shown.)

History group configuration example

Network requirements

Configure the RMON history group on the RMON agent in [Figure 30](#) to gather periodical traffic statistics for GigabitEthernet 1/0/1 every one minute.

Figure 30 Network diagram



Configuration procedure

Configure the RMON history group on the RMON agent to gather traffic statistics every one minute for GigabitEthernet 1/0/1. Retain up to eight records for the interface in the history statistics table.

```
<Sysname> system-view
[Sysname] interface gigabitethernet 1/0/1
[Sysname-GigabitEthernet1/0/1] rmon history 1 buckets 8 interval 60 owner user1
```

Display the history data collected for GigabitEthernet 1/0/1.

```
[Sysname-GigabitEthernet1/0/1] display rmon history
HistoryControlEntry 2 owned by null is VALID
  Samples interface      : GigabitEthernet1/0/1<ifIndex.3>
  Sampling interval      : 10(sec) with 8 buckets max
  Sampled values of record 1 :
    dropevents           : 0           , octets           : 834
    packets              : 8           , broadcast packets : 1
    multicast packets    : 6           , CRC alignment errors : 0
    undersize packets    : 0           , oversize packets  : 0
    fragments            : 0           , jabbers           : 0
    collisions           : 0           , utilization        : 0
  Sampled values of record 2 :
    dropevents           : 0           , octets           : 962
    packets              : 10          , broadcast packets : 3
    multicast packets    : 6           , CRC alignment errors : 0
    undersize packets    : 0           , oversize packets  : 0
    fragments            : 0           , jabbers           : 0
    collisions           : 0           , utilization        : 0
  Sampled values of record 3 :
    dropevents           : 0           , octets           : 830
    packets              : 8           , broadcast packets : 0
    multicast packets    : 6           , CRC alignment errors : 0
    undersize packets    : 0           , oversize packets  : 0
    fragments            : 0           , jabbers           : 0
    collisions           : 0           , utilization        : 0
  Sampled values of record 4 :
    dropevents           : 0           , octets           : 933
    packets              : 8           , broadcast packets : 0
    multicast packets    : 7           , CRC alignment errors : 0
    undersize packets    : 0           , oversize packets  : 0
```

```

    fragments      : 0          , jabbers          : 0
    collisions     : 0          , utilization       : 0
Sampled values of record 5 :
    dropevents    : 0          , octets           : 898
    packets       : 9          , broadcast packets : 2
    multicast packets : 6      , CRC alignment errors : 0
    undersize packets : 0      , oversize packets  : 0
    fragments     : 0          , jabbers          : 0
    collisions    : 0          , utilization       : 0
Sampled values of record 6 :
    dropevents    : 0          , octets           : 898
    packets       : 9          , broadcast packets : 2
    multicast packets : 6      , CRC alignment errors : 0
    undersize packets : 0      , oversize packets  : 0
    fragments     : 0          , jabbers          : 0
    collisions    : 0          , utilization       : 0
Sampled values of record 7 :
    dropevents    : 0          , octets           : 766
    packets       : 7          , broadcast packets : 0
    multicast packets : 6      , CRC alignment errors : 0
    undersize packets : 0      , oversize packets  : 0
    fragments     : 0          , jabbers          : 0
    collisions    : 0          , utilization       : 0
Sampled values of record 8 :
    dropevents    : 0          , octets           : 1154
    packets       : 13         , broadcast packets : 1
    multicast packets : 6      , CRC alignment errors : 0
    undersize packets : 0      , oversize packets  : 0
    fragments     : 0          , jabbers          : 0
    collisions    : 0          , utilization       : 0

```

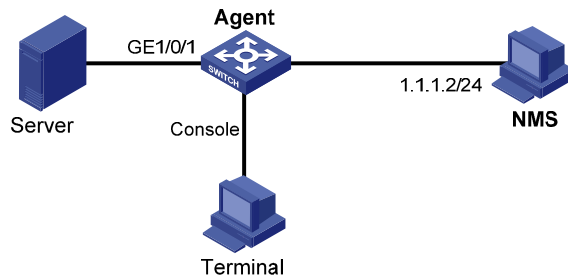
On the configuration terminal, get the traffic statistics through SNMP. (Details not shown.)

Alarm group configuration example

Network requirements

Configure the RMON alarm group on the RMON agent in [Figure 31](#) to send alarms in traps when the five-second incoming traffic statistic on GigabitEthernet 1/0/1 crosses the rising threshold or drops below the falling threshold.

Figure 31 Network diagram



Configuration procedure

Configure the SNMP agent with the same SNMP settings as the NMS at 1.1.1.2. This example uses SNMPv1, read community **public**, and write community **private**.

```
<Sysname> system-view
[Sysname] snmp-agent
[Sysname] snmp-agent community read public
[Sysname] snmp-agent community write private
[Sysname] snmp-agent sys-info version v1
[Sysname] snmp-agent trap enable
[Sysname] snmp-agent target-host trap address udp-domain 1.1.1.2 params securityname public
```

Configure the RMON statistics group to gather traffic statistics for GigabitEthernet 1/0/1.

```
[Sysname] interface gigabitethernet 1/0/1
[Sysname-GigabitEthernet1/0/1] rmon statistics 1 owner user1
[Sysname-GigabitEthernet1/0/1] quit
```

Create an RMON event entry and an RMON alarm entry so the RMON agent sends traps when the delta sampling value of node 1.3.6.1.2.1.16.1.1.1.4.1 exceeds 100 or drops below 50.

```
[Sysname] rmon event 1 trap public owner user1
[Sysname] rmon alarm 1 1.3.6.1.2.1.16.1.1.1.4.1 5 delta rising-threshold 100 1
falling-threshold 50 1
```

Display the RMON alarm entry configuration.

```
<Sysname> display rmon alarm 1
AlarmEntry 1 owned by null is Valid.
  Samples type          : delta
  Variable formula      : 1.3.6.1.2.1.16.1.1.1.4.1<etherStatsOctets.1>
  Sampling interval     : 5(sec)
  Rising threshold      : 100(linked with event 1)
  Falling threshold     : 50(linked with event 2)
  When startup enables  : risingOrFallingAlarm
  Latest value          : 0
```

Display statistics for GigabitEthernet 1/0/1.

```
<Sysname> display rmon statistics gigabitethernet 1/0/1
EtherStatsEntry 1 owned by user1-rmon is VALID.
  Interface : GigabitEthernet1/0/1<ifIndex.3>
  etherStatsOctets      : 57329      , etherStatsPkts          : 455
  etherStatsBroadcastPkts : 53      , etherStatsMulticastPkts : 353
```

```
etherStatsUndersizePkts : 0          , etherStatsOversizePkts : 0
etherStatsFragments     : 0          , etherStatsJabbers     : 0
etherStatsCRCAlignErrors : 0          , etherStatsCollisions  : 0
etherStatsDropEvents (insufficient resources): 0
Packets received according to length:
64      : 7          , 65-127 : 413          , 128-255 : 35
256-511: 0          , 512-1023: 0          , 1024-1518: 0
```

Query alarm events on the NMS. (Details not shown.)

On the RMON agent, alarm event messages are displayed when events occur. The following is a sample output:

```
[Sysname]
#Jan 27 16:31:34:12 2011 Sysname RMON/2/ALARMFALL:Trap 1.3.6.1.2.1.16.0.2 Alarm table 1
monitors 1.3.6.1.2.1.16.1.1.1.4.1 with sample type 2,has sampled alarm value 0 less than(or
=) 50.
```

Configuring port mirroring

Both bridge mode (Layer 2) and route mode (Layer 3) Ethernet ports support port mirroring. The term *interface* in this chapter collectively refers to these two types of ports. You can use the **port link-mode** command to set an Ethernet port to operate in bridge or route mode (see *Layer 2—LAN Switching Configuration Guide*).

Only the HP 5500 EI switch series supports configuring port mirroring on Layer 3 Ethernet interfaces.

Introduction to port mirroring

Port mirroring is the process of copying the packets passing through a port to the monitor port connecting to a monitoring device for packet analysis.

Terminologies of port mirroring

Mirroring source

The mirroring source can be one or more monitored ports. Packets (called "mirrored packets") passing through them are copied to a port connecting to a monitoring device for packet analysis. Such a port is called a "source port" and the device where the port resides is called a "source device".

Mirroring destination

The mirroring destination is the destination port (also known as the monitor port) of mirrored packets and connects to the data monitoring device. The device where the monitor port resides is called the "destination device." The monitor port forwards mirrored packets to its connected monitoring device.

A monitor port may receive multiple duplicates of a packet in some cases because it can monitor multiple mirroring sources. For example, assume that Port 1 is monitoring bidirectional traffic on Port 2 and Port 3 on the same device. If a packet travels from Port 2 to Port 3, two duplicates of the packet will be received on Port 1.

Mirroring direction

The mirroring direction indicates that the inbound, outbound, or bidirectional traffic can be copied on a mirroring source.

- Inbound: Copies packets received on a mirroring source.
- Outbound: Copies packets sent out of a mirroring source.
- Bidirectional: Copies packets both received and sent on a mirroring source.

Mirroring group

Port mirroring is implemented through mirroring groups, which fall into local, remote source, and remote destination mirroring groups. For more information about the mirroring groups, see "[Port mirroring classification and implementation](#)."

Reflector port, egress port, and remote probe VLAN

The reflector port, remote probe VLAN, and egress port are used for Layer 2 remote port mirroring. The remote probe VLAN specially transmits mirrored packets to the destination device. Both the reflector port and egress port reside on a source device and send mirrored packets to the remote probe VLAN. The

egress port must belong to the remote probe VLAN while the reflector port may not. For more information about the source device, destination device, reflector port, egress port, and remote probe VLAN, see "Port mirroring classification and implementation."

NOTE:

The reflector port is used to enable local mirroring to support multiple monitor ports.

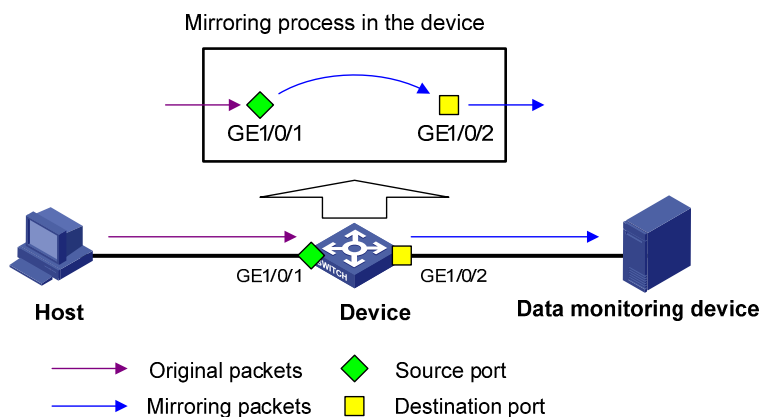
Port mirroring classification and implementation

According to the locations of the mirroring source and the mirroring destination, port mirroring falls into local port mirroring and remote port mirroring.

Local port mirroring

In local port mirroring, the mirroring source and the mirroring destination are on the same device. A mirroring group that contains the mirroring source and the mirroring destination on the device is called a "local mirroring group".

Figure 32 Local port mirroring implementation



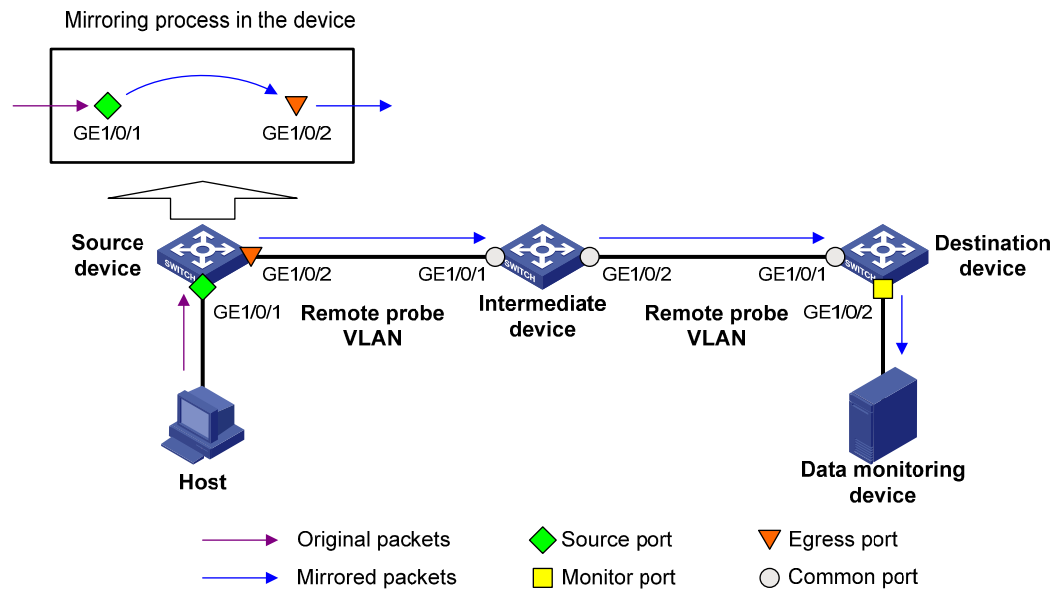
As shown in [Figure 32](#), the source port GigabitEthernet 1/0/1 and monitor port GigabitEthernet 1/0/2 reside on the same device. Packets of GigabitEthernet 1/0/1 are copied to GigabitEthernet 1/0/2, which then forwards the packets to the data monitoring device for analysis.

Remote port mirroring

In remote port mirroring, the mirroring source and the mirroring destination reside on different devices and in different mirroring groups. The mirroring group that contains the mirroring source or the mirroring destination is called a "remote source/destination group". The devices between the source devices and destination device are intermediate devices.

Because the source and destination devices are on the same Layer 2 network, remote port mirroring is also referred to Layer 2 remote port mirroring.

Figure 33 Layer 2 remote port mirroring implementation



On the network shown in [Figure 33](#),

The source device does the following:

1. Copies the packets received on the source port GigabitEthernet 1/0/1 to the egress port GigabitEthernet 1/0/2.
2. Forwards the packets to the intermediate device, which then broadcasts the packets in the remote probe VLAN.
3. Transmits the packets to the destination device via the intermediate device.

Then, the destination device does the following:

4. Receives the mirrored packets.
5. Compares their VLAN IDs to the ID of the remote probe VLAN configured in the remote destination group.
6. If the VLAN IDs of these mirrored packets match the remote probe VLAN ID, the device forwards them to the data monitoring device through the monitor port GigabitEthernet 1/0/2.

Allow remote probe VLAN to pass through the intermediate devices to make sure the source device and the destination device can communicate at Layer 2 in the remote probe VLAN.

For a mirrored packet to successfully arrive at the remote destination device, make sure the VLAN ID of the mirrored packet is not removed or changed. Otherwise, the Layer 2 remote port mirroring configuration will fail.

To monitor both the received and sent packets of a port in a mirroring group, you must use the **mac-address mac-learning disable** command on the source, intermediate, and destination devices to disable MAC address learning of the remote probe VLAN. For more information about the **mac-address mac-learning disable** command, see *Layer 2—LAN Switch Command Reference*.

Configuring local port mirroring

Local port mirroring configuration task list

Configure a local mirroring group and then configure one or more source ports and a monitor port for the local mirroring group.

Complete these tasks to configure local port mirroring:

| Task | Remarks |
|---|----------|
| Creating a local mirroring group | Required |
| Configuring source ports for the local mirroring group | Required |
| Configuring the monitor port for the local mirroring group | Required |
| Using the remote probe VLAN to enable local mirroring to support multiple monitor ports | Optional |

Creating a local mirroring group

| Step | Command | Remarks |
|------------------------------------|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Create a local mirroring group. | mirroring-group <i>group-id</i> local | No local mirroring group exists by default. |

NOTE:

A local mirroring group takes effect only after you configure a monitor port and source ports for it.

Configuring source ports for the local mirroring group

If you use system view, you can use a list to configure multiple source ports for a mirroring group at one time. If you use interface view, you can assign only the current port to the group as a source port, so you must repeat the step for each additional port.

Configuration restrictions and guidelines

- A mirroring group can contain multiple source ports.
- A port can belong to only one mirroring group.

Configuring source ports in system view

| Step | Command | Remarks |
|----------------------------|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Configure source ports. | mirroring-group <i>group-id</i>
mirroring-port <i>mirroring-port-list</i>
{ both inbound outbound } | By default, no source port is configured for a local mirroring group. |

Configuring a source port in interface view

| Step | Command | Remarks |
|---|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter interface view. | interface <i>interface-type</i>
<i>interface-number</i> | N/A |
| 3. Configure the current port as a source port. | [mirroring-group <i>group-id</i>]
mirroring-port { both inbound outbound } | By default, a port does not serve as a source port for any local mirroring group. |

Configuring the monitor port for the local mirroring group

You can configure the monitor port for a mirroring group in system view, or assign the current port to a mirroring group as the monitor port in interface view. The two methods lead to the same result.

Configuration restrictions and guidelines

- A mirroring group contains only one monitor port.
- To make sure that the mirroring function works properly, do not assign the monitor port to a source VLAN, or enable the spanning tree feature on the monitor port.
- HP recommends you use a monitor port for port mirroring only. This is to make sure that the data monitoring device receives and analyzes only the mirrored traffic rather than a mix of mirrored traffic and normally forwarded traffic.
- You cannot configure the monitor port in a mirroring group as a port in a RRPP ring.

Configuring the monitor port in system view

| Step | Command | Remarks |
|--------------------------------|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Configure the monitor port. | mirroring-group <i>group-id</i>
monitor-port <i>monitor-port-id</i> | By default, no monitor port is configured for a local mirroring group. |

Configuring the monitor port in interface view

| Step | Command | Remarks |
|--|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter interface view. | interface <i>interface-type</i>
<i>interface-number</i> | N/A |
| 3. Configure the current port as the monitor port. | [mirroring-group <i>group-id</i>]
monitor-port | By default, a port does not serve as the monitor port for any local mirroring group. |

Using the remote probe VLAN to enable local mirroring to support multiple monitor ports

In typical local port mirroring configuration, you can configure only one monitor port in a local mirroring group. As a result, you cannot monitor traffic of a local device on multiple data monitoring devices. To do that, you can take advantage of the remote probe VLAN used in Layer 2 remote mirroring.

In Layer 2 remote port mirroring, a remote probe VLAN is configured, and the mirrored packets are broadcast within the remote probe VLAN. By connecting multiple data monitoring devices to the member ports of the remote probe VLAN, you can monitor the traffic of the local device on multiple data monitoring devices.

Configure this feature in the following steps:

1. Configure a remote source mirroring group on the local device
2. Configure the monitored ports on the device as source ports of this mirroring group
3. Configure a remote probe VLAN for this mirroring group
4. Assign the ports connecting the data monitoring devices to the remote probe VLAN

In this way, when packets mirrored on the monitored ports are broadcast in the remote probe VLAN, they will be sent out of the ports connecting the data monitoring devices, and all the data monitoring devices can thus receive these mirrored packets.

Configuration restrictions and guidelines

- The reflector port of a remote source mirroring group must be an access port and belong to the default VLAN, VLAN 1.
- HP recommends that you configure an unused port as the reflector port of a remote source mirroring group and disable STP on it.
- A mirroring group can contain multiple source ports.
- To make sure that the port mirroring function works properly, do not assign a source port to the remote probe VLAN.
- If you have already configured a reflector port for a remote source mirroring group, you can no longer configure an egress port for it.
- A VLAN can serve as the remote probe VLAN for only one remote source mirroring group. HP recommends you use the remote probe VLAN for port mirroring exclusively. Do not create a VLAN interface for the VLAN or configure any other features for the VLAN.
- A remote probe VLAN must be a static VLAN. To remove the VLAN configured as a remote probe VLAN, you must first remove the remote probe VLAN with the **undo mirroring-group remote-probe vlan** command.
- If the remote probe VLAN of a remote mirroring group is removed, the remote mirroring group will become invalid.
- The link type of monitor ports configured for port mirroring must be access.

Configuration procedure

To configure local port mirroring with multiple monitor ports:

| Step | Command | Remarks |
|-----------------------|--------------------|---------|
| 1. Enter system view. | system-view | N/A |

| Step | Command | Remarks |
|---|---|---|
| 2. Create a remote source mirroring group. | mirroring-group <i>group-id</i>
remote-source | By default, no mirroring group exists on a device. |
| 3. Configure source ports for the remote source mirroring group. | <ul style="list-style-type: none"> • (Approach 1) In system view:
mirroring-group <i>group-id</i>
mirroring-port <i>mirroring-port-list</i>
{ both inbound outbound } • (Approach 2) In interface view: <ul style="list-style-type: none"> a. interface <i>interface-type</i>
<i>interface-number</i> b. [mirroring-group <i>group-id</i>]
mirroring-port { both inbound outbound } c. quit | Use either approach.
By default, no source port is configured for a mirroring group. |
| 4. Configure the reflector port for the remote source mirroring group. | mirroring-group <i>group-id</i> reflector-port
<i>reflector-port</i> | By default, no reflector port is configured for a mirroring group. |
| 5. Create the remote probe VLAN and enter VLAN view. | vlan <i>vlan-id</i> | By default, no remote probe VLAN is configured for a mirroring group. |
| 6. Assign monitor ports to the remote probe VLAN. | port <i>interface-list</i> | By default, a newly-created VLAN does not have any member port. |
| 7. Return to system view. | quit | N/A |
| 8. Configure the remote probe VLAN for the remote source mirroring group. | mirroring-group <i>group-id</i>
remote-probe vlan <i>rprobe-vlan-id</i> | By default, no remote probe VLAN is configured for a mirroring group. |

Configuring Layer 2 remote port mirroring

Layer 2 remote port mirroring configuration task list

Configuring Layer 2 remote port mirroring is to configure remote mirroring groups. To do that, configure the remote source group on the source device and configure the cooperating remote destination group on the destination device. If an intermediate device exists, allow the remote probe VLAN to pass through the intermediate device.

NOTE:

HP recommends you not enable GARP VLAN Registration Protocol (GVRP). If GVRP is enabled, GVRP may register the remote probe VLAN to unexpected ports, resulting in undesired duplicates. For more information about GVRP, see *Layer 2—LAN Switching Configuration Guide*.

Configure the following on the source device:

- Source ports
- Remote probe VLAN
- The egress port

Then, configure the following on the destination device:

- Remote probe VLAN
- Monitor port

Complete these tasks to configure Layer 2 remote port mirroring:

| Task | Remarks | |
|--|--|----------|
| Configuring a remote source group | Creating a remote source group | Required |
| | Configuring source ports for the remote source group | Required |
| | Configuring the egress port for the remote source group | Required |
| | Configuring the remote probe VLAN for the remote source group | Required |
| Configuring a remote destination group | Creating a remote destination group | Required |
| | Configuring the monitor port for the remote destination group | Required |
| | Configuring the remote probe VLAN for the remote destination group | Required |
| | Assigning the monitor port to the remote probe VLAN | Required |

Configuring a remote source group (on the source device)

Creating a remote source group

| Step | Command | Remarks |
|----------------------------------|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Create a remote source group. | mirroring-group <i>group-id</i> remote-source | By default, no remote source group exists on a device. |

Configuring source ports for the remote source group

If you use system view, you can use a list to configure multiple source ports for a mirroring group at one time. If you use interface view, you can assign only the current port to the group as a source port, so you must repeat the step for each additional port.

1. Configuration restrictions and guidelines:
 - A mirroring group can contain multiple source ports.
 - A port can belong to only one mirroring group.

2. Configuration procedure:

To configure source ports for the remote source group in system view:

| Step | Command | Remarks |
|--|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Configure source ports for the remote source group. | mirroring-group <i>group-id</i>
mirroring-port <i>mirroring-port-list</i>
{ both inbound outbound } | By default, no source port is configured for a remote source group. |

To configure a source port for the remote source group in interface view:

| Step | Command | Remarks |
|---|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter interface view. | interface <i>interface-type</i>
<i>interface-number</i> | N/A |
| 3. Configure the current port as a source port. | [mirroring-group <i>group-id</i>]
mirroring-port { both inbound outbound } | By default, a port does not serve as a source port for any remote source group. |

Configuring the egress port for the remote source group

You can configure the egress port for a mirroring group in system view, or assign the current port to it as the egress port in interface view. The two configuration modes lead to the same result.

To make sure that the mirroring function works properly, disable these functions on the egress port: the spanning tree feature, 802.1X, IGMP snooping, static ARP, and MAC address learning.

To configure the egress port for the remote source group in system view:

| Step | Command | Remarks |
|---|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Configure the egress port for the remote source group. | mirroring-group <i>group-id</i>
monitor-egress
<i>monitor-egress-port</i> | By default, no egress port is configured for a remote source group. |

To configure the egress port for the remote source group in interface view:

| Step | Command | Remarks |
|---|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter interface view. | interface <i>interface-type</i>
<i>interface-number</i> | N/A |
| 3. Configure the current port as the egress port. | mirroring-group <i>group-id</i>
monitor-egress | By default, a port does not serve as the egress port for any remote source group. |

NOTE:

- A mirroring group contains only one egress port.
- A source port of an existing mirroring group cannot be configured as an egress port.

Configuring the remote probe VLAN for the remote source group

Before configuring a remote probe VLAN, create a static VLAN that will serve as the remote probe VLAN for the remote source group.

1. Configuration restrictions and guidelines:
 - A VLAN can serve for only one mirroring group.
 - When a VLAN is configured as a remote probe VLAN, you must remove the remote probe VLAN configuration before deleting the VLAN.
 - When you remove the configuration of a remote probe VLAN, an active mirroring group becomes inactive.

- When a VLAN is configured as a remote probe VLAN, use the remote probe VLAN for port mirroring exclusively. Do not create a VLAN interface for the VLAN or configure any other features for the VLAN.
 - The remote mirroring groups on the source device and destination device must use the same remote probe VLAN.
2. Configuration procedure:
To configure the remote probe VLAN for the remote source group:

| Step | Command | Remarks |
|-------------------------------------|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Configure the remote probe VLAN. | mirroring-group <i>group-id</i>
remote-probe vlan
<i>rprobe-vlan-id</i> | By default, no remote probe VLAN is configured for a remote source group. |

Configuring a remote destination group (on the destination device)

To configure a remote destination group, make the following configurations on the destination device:

Creating a remote destination group

| Step | Command | Remarks |
|---------------------------------------|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Create a remote destination group. | mirroring-group <i>group-id</i>
remote-destination | By default, no remote destination group exists on a device. |

Configuring the monitor port for the remote destination group

You can configure the monitor port for a mirroring group in system view, or assign the current port to a mirroring group as the monitor port in interface view. The two methods lead to the same result.

1. Configuration restrictions and guidelines:
- A mirroring group contains only one monitor port.
 - To make sure that the mirroring function works properly, do not enable the spanning tree feature on the monitor port.
 - HP recommends you use a monitor port only for port mirroring. This is to make sure that the data monitoring device receives and analyzes only the mirrored traffic rather than a mix of mirrored traffic and normally forwarded traffic.
 - You cannot configure the monitor port in a mirroring group as a port in a RRPP ring.
2. Configuration procedure:

To configure the monitor port for the remote destination group in system view:

| Step | Command | Remarks |
|-----------------------|--------------------|---------|
| 1. Enter system view. | system-view | N/A |

| Step | Command | Remarks |
|--------------------------------|--|---|
| 2. Configure the monitor port. | mirroring-group <i>group-id</i>
monitor-port <i>monitor-port-id</i> | By default, no monitor port is configured for a remote destination group. |

To configure the monitor port for the remote destination group in interface view:

| Step | Command | Remarks |
|--|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter interface view. | interface <i>interface-type</i>
<i>interface-number</i> | N/A |
| 3. Configure the current port as the monitor port. | [mirroring-group <i>group-id</i>]
monitor-port | By default, a port does not serve as the monitor port for any remote destination group. |

Configuring the remote probe VLAN for the remote destination group

- Configuration restrictions and guidelines:
 - A VLAN can serve for only one mirroring group.
 - When a VLAN is configured as a remote probe VLAN, use the remote probe VLAN for port mirroring exclusively. Do not configure a VLAN interface for the VLAN or configure any other features for the VLAN.
 - When a VLAN is configured as a remote probe VLAN, you must remove the remote probe VLAN configuration before deleting the VLAN.
 - When you remove the configuration of a remote probe VLAN, an active mirroring group becomes inactive.
- Configuration procedure:

| Step | Command | Remarks |
|-------------------------------------|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Configure the remote probe VLAN. | mirroring-group <i>group-id</i>
remote-probe vlan <i>rprobe-vlan-id</i> | By default, no remote probe VLAN is configured for a remote destination group. |

Assigning the monitor port to the remote probe VLAN

| Step | Command | Remarks |
|--|---|-------------------------|
| 1. Enter system view. | system-view | N/A |
| 2. Enter the interface view of the monitor port. | interface <i>interface-type</i> <i>interface-number</i> | N/A |
| 3. Assign the port to the probe VLAN. | <ul style="list-style-type: none"> For an access port:
port access vlan <i>vlan-id</i> For a trunk port:
port trunk permit vlan <i>vlan-id</i> For a hybrid port:
port hybrid vlan <i>vlan-id</i> { tagged untagged } | Use one of the commands |

For more information about the **port access vlan**, **port trunk permit vlan**, and **port hybrid vlan** commands, see *Layer 2—LAN Switching Command Reference*.

Displaying and maintaining port mirroring

| Task | Command | Remarks |
|--|--|-----------------------|
| Display the configuration of mirroring groups. | display mirroring-group { <i>group-id</i> all local remote-destination remote-source } [{ begin exclude include } <i>regular-expression</i>] | Available in any view |

Port mirroring configuration examples

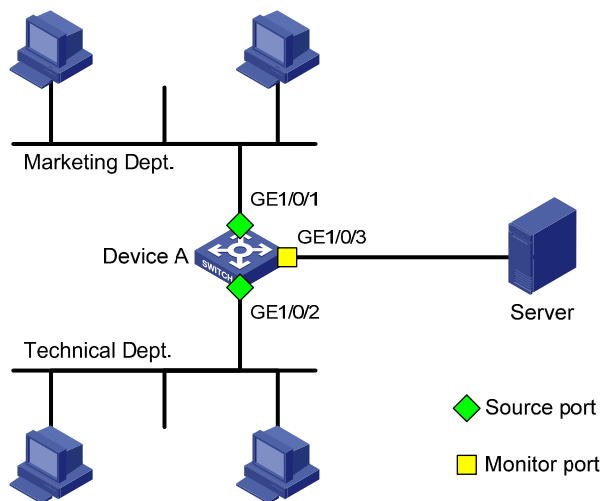
Local port mirroring configuration example

Network requirements

On the network shown in [Figure 34](#):

- Device A connects to the marketing department through GigabitEthernet 1/0/1 and to the technical department through GigabitEthernet 1/0/2. It connects to the server through GigabitEthernet 1/0/3.
- Configure local port mirroring in source port mode to enable the server to monitor the bidirectional traffic of the marketing department and the technical department.

Figure 34 Network diagram



Configuration procedure

1. Create a local mirroring group:

```
# Create local mirroring group 1.
```

```
<DeviceA> system-view
```

```
[DeviceA] mirroring-group 1 local
```

```
# Configure GigabitEthernet 1/0/1 and GigabitEthernet 1/0/2 as source ports and port GigabitEthernet 1/0/3 as the monitor port.
```

```
[DeviceA] mirroring-group 1 mirroring-port GigabitEthernet 1/0/1 GigabitEthernet
1/0/2 both
[DeviceA] mirroring-group 1 monitor-port GigabitEthernet 1/0/3
# Disable the spanning tree feature on the monitor port GigabitEthernet 1/0/3.
[DeviceA] interface GigabitEthernet 1/0/3
[DeviceA-GigabitEthernet1/0/3] undo stp enable
[DeviceA-GigabitEthernet1/0/3] quit
```

2. Verify the configurations:

Display the configuration of all mirroring groups.

```
[DeviceA] display mirroring-group all
mirroring-group 1:
  type: local
  status: active
  mirroring port:
    GigabitEthernet1/0/1 both
    GigabitEthernet1/0/2 both
  monitor port: GigabitEthernet1/0/3
```

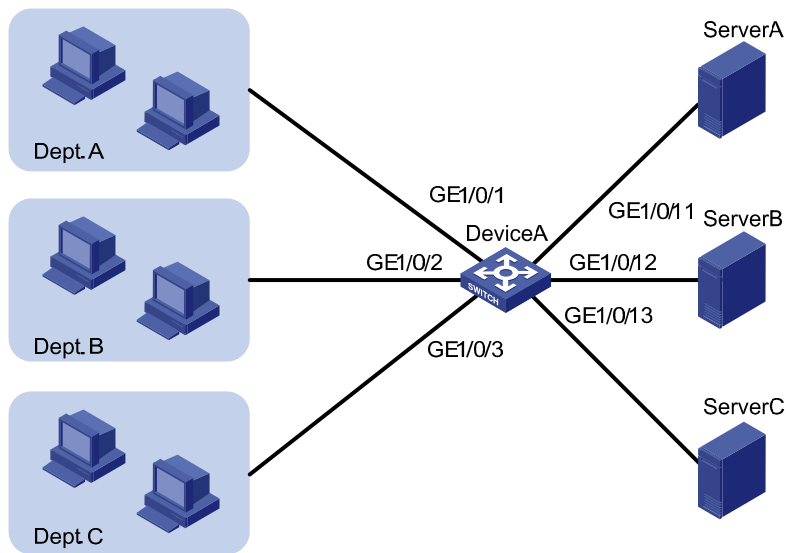
After the configurations are completed, you can monitor all the packets received and sent by the marketing department and the technical department on the server.

Local port mirroring with multiple monitor ports configuration example

Network requirements

As shown in [Figure 35](#), Dept. A, Dept. B, and Dept. C are connected to Device A through ports GigabitEthernet 1/0/1, GigabitEthernet 1/0/2, and GigabitEthernet 1/0/3, respectively. Configure port mirroring to enable all three data monitoring devices, Server A, Server B, and Server C, to monitor both the incoming and outgoing traffic of the three departments.

Figure 35 Network diagram



Configuration procedure

```
# Create remote source mirroring group 1.
<DeviceA> system-view
[DeviceA] mirroring-group 1 remote-source

# Configure GigabitEthernet 1/0/1 through GigabitEthernet 1/0/3 as source ports of remote source mirroring group 1.
[DeviceA] mirroring-group 1 mirroring-port gigabitethernet 1/0/1 to gigabitethernet 1/0/3 both

# Configure an unused port (GigabitEthernet 1/0/5 for example) of Device A as the reflector port of remote source mirroring group 1.
[DeviceA] mirroring-group 1 reflector-port GigabitEthernet 1/0/5

# Create VLAN 10 and assign the three ports (GigabitEthernet 1/0/11 through GigabitEthernet 1/0/13) connecting the three data monitoring devices to VLAN 10.
[DeviceA] vlan 10
[DeviceA-vlan10] port gigabitethernet 1/0/11 to gigabitethernet 1/0/13
[DeviceA-vlan10] quit

# Configure VLAN 10 as the remote probe VLAN of remote source mirroring group 1.
[DeviceA] mirroring-group 1 remote-probe vlan 10
```

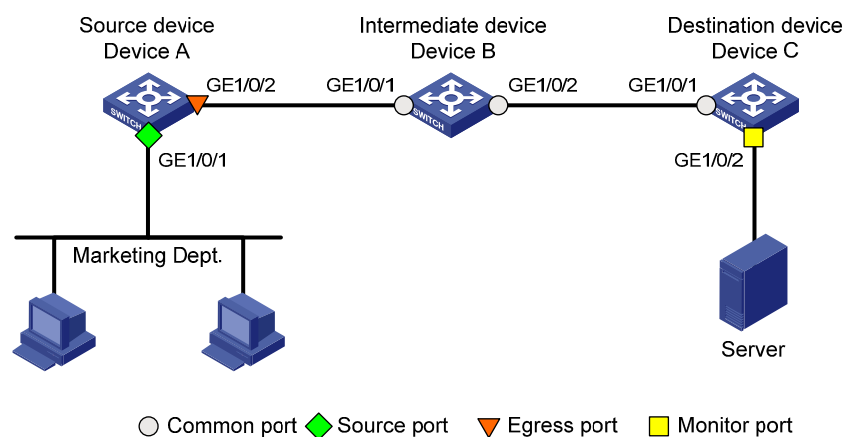
Layer 2 remote port mirroring configuration example

Network requirements

On the Layer 2 network shown in [Figure 36](#):

- Device A connects to the marketing department through GigabitEthernet 1/0/1 and connects to the trunk port GigabitEthernet 1/0/2 of Device B through the trunk port GigabitEthernet 1/0/1.
- Device C connects to the server through GigabitEthernet 1/0/2 and connects to the trunk port GigabitEthernet 1/0/1 of Device B through the trunk port GigabitEthernet 1/0/2.
- Configure Layer 2 remote port mirroring to enable the server to monitor the bidirectional traffic of the marketing department.

Figure 36 Network diagram



Configuration procedure

1. Configure Device A (the source device):

```

# Create a remote source group.
<DeviceA> system-view
[DeviceA] mirroring-group 1 remote-source
# Create VLAN 2 as the remote probe VLAN.
[DeviceA] vlan 2
# Disable MAC address learning for the remote probe VLAN.
[DeviceA-vlan2] mac-address mac-learning disable
[DeviceA-vlan2] quit
# Configure VLAN 2 as the remote probe VLAN of the mirroring group; configure GigabitEthernet
1/0/1 as a source port and GigabitEthernet 1/0/2 as the egress port in the mirroring group.
[DeviceA] mirroring-group 1 remote-probe vlan 2
[DeviceA] mirroring-group 1 mirroring-port GigabitEthernet 1/0/1 both
[DeviceA] mirroring-group 1 monitor-egress GigabitEthernet 1/0/2
# Configure output port GigabitEthernet 1/0/2 as a trunk port to permit the packets of VLAN 2 to
pass through, and disable the spanning tree feature on the port.
[DeviceA] interface GigabitEthernet 1/0/2
[DeviceA-GigabitEthernet1/0/2] port link-type trunk
[DeviceA-GigabitEthernet1/0/2] port trunk permit vlan 2
[DeviceA-GigabitEthernet1/0/2] undo stp enable
[DeviceA-GigabitEthernet1/0/2] quit

```

2. Configure Device B (the intermediate device):

```

# Create VLAN 2 as the remote probe VLAN.
<DeviceB> system-view
[DeviceB] vlan 2
# Disable MAC address learning for the remote probe VLAN.
[DeviceB-vlan2] mac-address mac-learning disable
[DeviceB-vlan2] quit
# Configure GigabitEthernet 1/0/1 as a trunk port that permits the packets of VLAN 2 to pass
through.
[DeviceB] interface GigabitEthernet 1/0/1
[DeviceB-GigabitEthernet1/0/1] port link-type trunk
[DeviceB-GigabitEthernet1/0/1] port trunk permit vlan 2
[DeviceB-GigabitEthernet1/0/1] quit
# Configure GigabitEthernet 1/0/2 as a trunk port that permits the packets of VLAN 2 to pass
through.
[DeviceB-GigabitEthernet1/0/1] quit
[DeviceB] interface GigabitEthernet 1/0/2
[DeviceB-GigabitEthernet1/0/2] port link-type trunk
[DeviceB-GigabitEthernet1/0/2] port trunk permit vlan 2
[DeviceB-GigabitEthernet1/0/2] quit

```

3. Configure Device C (the destination device):

```

# Configure GigabitEthernet 1/0/1 as a trunk port that permits the packets of VLAN 2 to pass
through.
<DeviceC> system-view
[DeviceC] interface GigabitEthernet 1/0/1
[DeviceC-GigabitEthernet1/0/1] port link-type trunk

```

```

[DeviceC-GigabitEthernet1/0/1] port trunk permit vlan 2
[DeviceC-GigabitEthernet1/0/1] quit
# Create a remote destination group.
[DeviceC] mirroring-group 1 remote-destination
# Create VLAN 2 as the remote probe VLAN.
[DeviceC] vlan 2
# Disable MAC address learning for the remote probe VLAN.
[DeviceC-vlan2] mac-address mac-learning disable
[DeviceC-vlan2] quit
# Configure VLAN 2 as the remote probe VLAN of the mirroring group and GigabitEthernet
1/0/2 as the monitor port of the mirroring group, disable the spanning tree feature on
GigabitEthernet 1/0/2, and assign the port to VLAN 2.
[DeviceC] mirroring-group 1 remote-probe vlan 2
[DeviceC] interface GigabitEthernet 1/0/2
[DeviceC-GigabitEthernet1/0/2] mirroring-group 1 monitor-port
[DeviceC-GigabitEthernet1/0/2] undo stp enable
[DeviceC-GigabitEthernet1/0/2] port access vlan 2
[DeviceC-GigabitEthernet1/0/2] quit

```

4. Verify the configurations.

After the configurations are completed, you can monitor all the packets received and sent by the marketing department on the server.

Configuring traffic mirroring

- Both bridge mode (Layer 2) and route mode (Layer 3) Ethernet ports support traffic mirroring. The term "interface" in this chapter collectively refers to these two types of ports. You can use the **port link-mode** command to set an Ethernet port to operate in bridge or route mode (see *Layer 2—LAN Switching Configuration Guide*).
- Only the HP 5500 EI switch series supports configuring traffic mirroring on Layer 3 Ethernet interfaces.

Introduction to traffic mirroring

Traffic mirroring copies the specified packets to the specified destination for packet analyzing and monitoring. It is implemented through QoS policies. In other words, you define traffic classes and configure match criteria to classify packets to be mirrored and then configure traffic behaviors to mirror packets that fit the match criteria to the specified destination. Traffic mirroring allows you to flexibly classify packets by defining match criteria and obtain accurate statistics.

You can configure the traffic to be mirrored to an interface, to a CPU, or to a VLAN.

- Mirroring traffic to an interface copies the matching packets to a destination interface.
- Mirroring traffic to a CPU copies the matching packets to a CPU.

For more information about QoS policies, traffic classes, and traffic behaviors, see *ACL and QoS Configuration Guide*.

Traffic mirroring configuration task list

| Task | Remarks |
|--|---|
| Configuring match criteria | Required |
| Configuring traffic mirroring of different types | Mirroring traffic to a port
Required |
| | Mirroring traffic to the CPU
Perform at least one configuration. |
| Configuring a QoS policy | Required |
| Applying a QoS policy | Apply a QoS policy to a port
Required |
| | Apply a QoS policy to a VLAN
Required |
| | Apply a QoS policy globally
Perform one of these configurations |
| | Apply a QoS policy to the control plane |

Configuring match criteria

| Step | Command | Remarks |
|-----------------------|--------------------|---------|
| 1. Enter system view. | system-view | N/A |

| Step | Command | Remarks |
|---|--|--|
| 2. Create a class and enter class view. | traffic classifier <i>tcl-name</i> [operator { and or }] | By default, no traffic class exists. |
| 3. Configure match criteria. | if-match <i>match-criteria</i> | By default, no match criterion is configured in a traffic class. |

For more information about the **traffic classifier** and **if-match** commands, see *ACL and QoS Command Reference*.

Configuring traffic mirroring of different types

In a traffic behavior, you can configure only one type of traffic mirroring.

Mirroring traffic to a port

| Step | Command | Remarks |
|---|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Create a behavior and enter behavior view. | traffic behavior
<i>behavior-name</i> | By default, no traffic behavior exists.
For more information about the traffic behavior command, see <i>ACL and QoS Command Reference</i> . |
| 3. Specify the destination interface for traffic mirroring. | mirror-to interface
<i>interface-type</i>
<i>interface-number</i> | By default, traffic mirroring is not configured in a traffic behavior.
You can specify up to four destination interfaces by executing the mirror-to interface command repeatedly. |

Mirroring traffic to the CPU

| Step | Command | Remarks |
|---|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Create a behavior and enter behavior view. | traffic behavior
<i>behavior-name</i> | By default, no traffic behavior exists.
For more information about the traffic behavior command, see <i>ACL and QoS Command Reference</i> . |
| 3. Mirror traffic to the CPU. | mirror-to cpu | By default, no traffic mirroring is configured in a traffic behavior. |

NOTE:

The CPU refers to the CPU of the device where ports with traffic mirroring configured reside.

Configuring a QoS policy

| Step | Command | Remarks |
|---|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Create a policy and enter policy view. | qos policy <i>policy-name</i> | By default, no policy exists. |
| 3. Associate a class with a traffic behavior in the QoS policy. | classifier <i>tcl-name</i> behavior <i>behavior-name</i> | By default, no traffic behavior is associated with a class. |

For more information about the **qos policy** and **classifier behavior** commands, see *ACL and QoS Command Reference*.

Applying a QoS policy

For more information about applying a QoS policy, see *ACL and QoS Configuration Guide*.

Apply a QoS policy to a port

By applying a QoS policy to an interface, you can mirror the traffic in a specified direction on the interface. A policy can be applied to multiple interfaces, but in one direction (inbound or outbound) of an interface, only one policy can be applied.

To apply a QoS policy to a port:

| Step | Command | Remarks |
|--|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter interface view or port group view. | <ul style="list-style-type: none"> Enter interface view:
interface <i>interface-type</i> <i>interface-number</i> Enter port group view:
port-group manual <i>port-group-name</i> | Use either command
Settings in interface view take effect on the current interface; settings in port group view take effect on all ports in the port group. |
| 3. Apply a policy to the interface, all ports in the port group, or the PVC. | qos apply policy <i>policy-name</i> { inbound outbound } | For more information about the qos apply policy command, see <i>ACL and QoS Command Reference</i> . |

Apply a QoS policy to a VLAN

You can apply a QoS policy to a VLAN to mirror the traffic in a specified direction on all ports in the VLAN.

To apply the QoS policy to a VLAN:

| Step | Command |
|----------------------------------|--|
| 1. Enter system view. | system-view |
| 2. Apply a QoS policy to a VLAN. | qos vlan-policy <i>policy-name</i> vlan <i>vlan-id-list</i> { inbound outbound } |

For more information about the **qos vlan-policy** command, see *ACL and QoS Command Reference*.

Apply a QoS policy globally

You can apply a QoS policy globally to mirror the traffic in a specified direction on all ports.

To apply a QoS policy globally:

| Step | Command |
|---------------------------------|---|
| 1. Enter system view. | system-view |
| 2. Apply a QoS policy globally. | qos apply policy <i>policy-name</i> global { inbound outbound } |

For more information about the **qos apply policy** command, see *ACL and QoS Command Reference*.

Apply a QoS policy to the control plane

You can apply a QoS policy to the control plane to mirror the traffic in the inbound direction of the control plane.

To apply a QoS policy to the control plane:

| Step | Command |
|---|---|
| 1. Enter system view. | system-view |
| 2. Enter control plane view. | control-plane slot <i>slot-number</i> |
| 3. Apply a QoS policy to the control plane. | qos apply policy <i>policy-name</i> inbound |

For more information about the **control-plane** and **qos apply policy** commands, see *ACL and QoS Command Reference*.

Displaying and maintaining traffic mirroring

| Task | Command | Remarks |
|--|---|-----------------------|
| Display user-defined traffic behavior configuration information. | display traffic behavior user-defined
[<i>behavior-name</i>] [[{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display user-defined QoS policy configuration information. | display qos policy user-defined [<i>policy-name</i>]
[classifier <i>tcl-name</i>] [[{ begin exclude include } <i>regular-expression</i>] | Available in any view |

For more information about the **display traffic behavior** and **display qos policy** commands, see *ACL and QoS Command Reference*.

Traffic mirroring configuration example

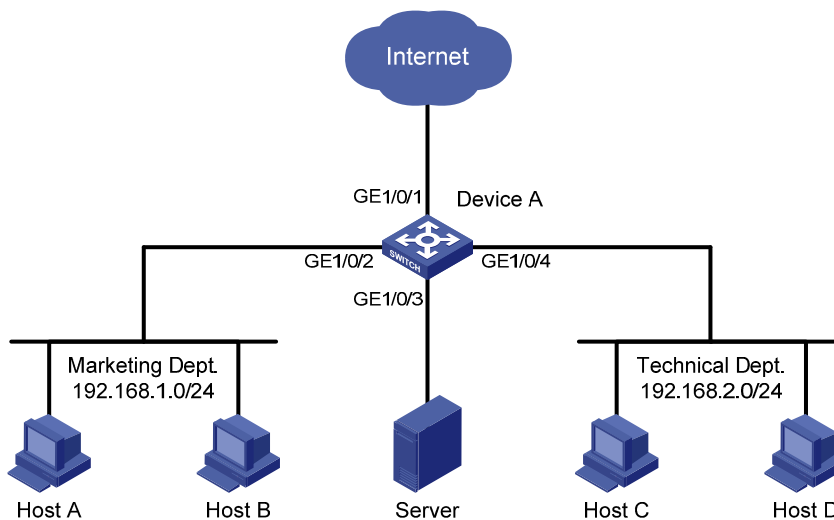
Traffic mirroring configuration example

Network requirements

As shown in Figure 37:

- Different departments of a company use IP addresses on different subnets. The marketing and technology departments use the IP addresses on subnets 192.168.1.0/24 and 192.168.2.0/24 respectively. The working hour of the company is from 8:00 to 18:00 on weekdays.
- Configure traffic mirroring so that the server can monitor the traffic that the technology department sends to access the Internet, and IP traffic that the technology department sends to the marketing department.

Figure 37 Network diagram



Configuration procedure

1. Monitor the traffic sent by the technology department to access the Internet:

Create ACL 3000 to allow packets from the technology department (on subnet 192.168.2.0/24) to access the Internet.

```
<DeviceA> system-view
[DeviceA] acl number 3000
[DeviceA-acl-adv-3000] rule permit tcp source 192.168.2.0 0.0.0.255 destination-port eq www
[DeviceA-acl-adv-3000] quit
```

Create traffic class **tech_c**, and configure the match criterion as ACL 3000.

```
[DeviceA] traffic classifier tech_c
[DeviceA-classifier-tech_c] if-match acl 3000
[DeviceA-classifier-tech_c] quit
```

Create traffic behavior **tech_b**, and configure the action of mirroring traffic to port GigabitEthernet 1/0/3.

```
[DeviceA] traffic behavior tech_b
[DeviceA-behavior-tech_b] mirror-to interface GigabitEthernet 1/0/3
```

```
[DeviceA-behavior-tech_b] quit
# Create QoS policy tech_p, and associate traffic class tech_c with traffic behavior tech_b in the
QoS policy.
[DeviceA] qos policy tech_p
[DeviceA-qospolicy-tech_p] classifier tech_c behavior tech_b
[DeviceA-qospolicy-tech_p] quit
# Apply QoS policy tech_p to the outgoing packets of GigabitEthernet 1/0/1.
[DeviceA] interface GigabitEthernet 1/0/1
[DeviceA-GigabitEthernet1/0/1] qos apply policy tech_p outbound
[DeviceA-GigabitEthernet1/0/1] quit
```

2. Monitor the traffic that the technology department sends to the marketing department:

```
# Configure a time range named work to cover the time from 8: 00 to 18: 00 in working days.
[DeviceA] time-range work 8:0 to 18:0 working-day
# Create ACL 3001 to allow packets sent from the technology department (on subnet
192.168.2.0/24) to the marketing department (on subnet 192.168.1.0/24).
[DeviceA] acl number 3001
[DeviceA-acl-adv-3001] rule permit ip source 192.168.2.0 0.0.0.255 destination
192.168.1.0 0.0.0.255 time-range work
[DeviceA-acl-adv-3001] quit
# Create traffic class mkt_c, and configure the match criterion as ACL 3001.
[DeviceA] traffic classifier mkt_c
[DeviceA-classifier-mkt_c] if-match acl 3001
[DeviceA-classifier-mkt_c] quit
# Create traffic behavior mkt_b, and configure the action of mirroring traffic to port
GigabitEthernet 1/0/3.
[DeviceA] traffic behavior mkt_b
[DeviceA-behavior-mkt_b] mirror-to interface GigabitEthernet 1/0/3
[DeviceA-behavior-mkt_b] quit
# Create QoS policy mkt_p, and associate traffic class mkt_c with traffic behavior mkt_b in the
QoS policy.
[DeviceA] qos policy mkt_p
[DeviceA-qospolicy-mkt_p] classifier mkt_c behavior mkt_b
[DeviceA-qospolicy-mkt_p] quit
# Apply QoS policy mkt_p to the outgoing packets of GigabitEthernet 1/0/2.
[DeviceA] interface GigabitEthernet 1/0/2
[DeviceA-GigabitEthernet1/0/2] qos apply policy mkt_p outbound
```

3. Verify the configurations.

After completing the configurations, through the server, you can monitor all traffic sent by the technology department to access the Internet and the IP traffic that the technology department sends to the marketing department during working hours.

Configuring NQA

Overview

Network Quality Analyzer (NQA) can perform various types of tests and collect network performance and service quality parameters such as delay jitter, time for establishing a TCP connection, time for establishing an FTP connection, and file transfer rate.

With the NQA test results, you can diagnose and locate network faults, be aware of network performance in time and take proper actions to correct any problems.

NQA features

Supporting multiple test types

Ping uses only the Internet Control Message Protocol (ICMP) to test the reachability of the destination host and the round-trip time. As an enhancement to ping, NQA supports more test types and functions.

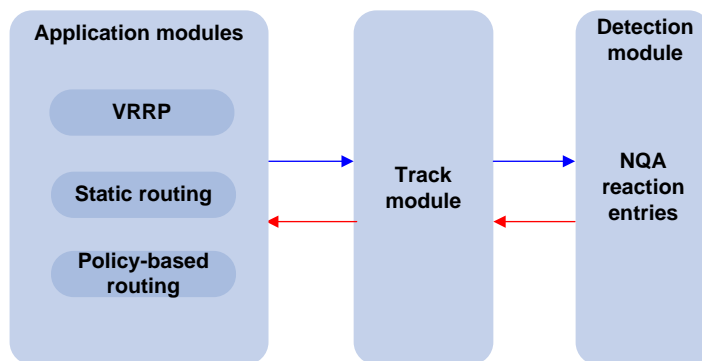
NQA supports 11 test types: ICMP echo, DHCP, DNS, FTP, HTTP, UDP jitter, SNMP, TCP, UDP echo, voice, and DLSw.

NQA enables the client to send probe packets of different test types to detect the protocol availability and response time of the peer. Test results help you understand network performance.

Supporting the collaboration function

Collaboration is implemented by establishing reaction entries to monitor the detection results of NQA probes. If the number of consecutive probe failures reaches a limit, NQA informs the track module of the detection result, and the track module triggers other application modules to take predefined.

Figure 38 Implement collaboration



The collaboration comprises the following parts: the application modules, the track module, and the detection modules.

- A detection module monitors objects, such as the link status, and network performance, and informs the track module of detection results.
- Upon the detection results, the track module changes the status of the track entry and informs the associated application module. The track module works between the application modules and the detection modules. It hides the differences among detection modules from application modules.

- The application module takes actions when the tracked object changes its state.

The following describes how a static route is monitored through collaboration.

1. NQA monitors the reachability to 192.168.0.88.
2. When 192.168.0.88 becomes unreachable, NQA notifies the track module of the change.
3. The track module notifies the state change to the static routing module
4. The static routing module sets the static route as invalid.

For more information about collaboration and the track module, see *High Availability Configuration Guide*.

Supporting threshold monitoring

NQA supports threshold monitoring for performance parameters such as average delay jitter and packet round-trip time. The performance parameters to be monitored are monitored elements. NQA monitors threshold violations for a monitored element, and reacts to certain measurement conditions (for example, sending trap messages to the network management server). This helps network administrators understand the network service quality and network performance.

- Monitored elements

[Table 8](#) describes the monitored elements and the NQA test types in which the elements can be monitored.

Table 8 Monitored elements and NQA test types

| Monitored elements | Test type supported |
|---|--|
| Probe duration | Tests excluding UDP jitter test and voice test |
| Count of probe failures | Tests excluding UDP jitter test and voice test |
| Packet round-trip time | UDP jitter test and voice test |
| Count of discarded packets | UDP jitter test and voice test |
| One-way delay jitter (source-to-destination and destination-to-source) | UDP jitter test and voice test |
| One-way delay (source-to-destination and destination-to-source) | UDP jitter test and voice test |
| Calculated Planning Impairment Factor (ICPIF) (see " Configuring voice tests ") | Voice test |
| Mean Opinion Scores (MOS) (see " Configuring voice tests ") | Voice test |

- Threshold types

The following threshold types are supported:

- **average**—Monitors the average value of monitored data in a test. If the average value in a test exceeds the upper threshold or goes below the lower threshold, a threshold violation occurs. For example, you can monitor the average probe duration in a test.
- **accumulate**—Monitors total number of times the monitored data violates the threshold in a test. If the total number of times reaches or exceeds a specific value, a threshold violation occurs.
- **consecutive**—Monitors the number of consecutive times the monitored data violates the threshold since the test group starts. If the monitored data violates the threshold consecutively for a specific number of times, a threshold violation occurs.

The counting for the average or accumulate threshold type is performed per test, but the counting for the consecutive type is performed after the test group starts.

- Triggered actions

The following actions may be triggered:

- **none**—NQA only records events for terminal display; it does not send trap information to the network management server. NQA DNS tests do not support the action of sending trap messages. The action to be triggered in DNS tests can only be the default one, **none**.
- **trap-only**—NQA records events and sends trap messages to the network management server.

- Reaction entry

In a reaction entry, a monitored element, a threshold type, and the action to be triggered are configured to implement threshold monitoring.

The state of a reaction entry can be invalid, over-threshold, or below-threshold, using the following workflow:

- Before an NQA test group starts, the reaction entry is in the state of invalid.
- After each test or probe, threshold violations are counted according to the threshold type and range configured in the entry. If the threshold is violated consecutively or accumulatively for a specific number of times, the state of the entry is set to over-threshold; otherwise, the state of the entry is set to below-threshold.

If the action to be triggered is configured as **trap-only** for a reaction entry, when the state of the entry changes, a trap message is generated and sent to the network management server.

NQA concepts

Test group

An NQA test group specifies test parameters including the test type, destination address, and destination port. Each test group is uniquely identified by an administrator name and operation tag. You can configure and schedule multiple NQA test groups to test different objects.

Test and probe

After the NQA test group starts, tests are performed at a specific interval. During each test, a specific number of probe operations are performed. Both the test interval and the number of probe operations per test are configurable. But only one probe operation is performed during one voice test.

In different test types, probe operation has the following different meanings:

- During a TCP or DLSw test, one probe operation means setting up one connection.
- During a UDP jitter or a voice test, one probe operation means continuously sending a specific number of probe packets. The number of probe packets is configurable.
- During an FTP, HTTP, DHCP, or DNS test, one probe operation means uploading or downloading a file, obtaining a web page, obtaining an IP address through DHCP, or translating a domain name to an IP address.
- During an ICMP echo or UDP echo test, one probe operation means sending an ICMP echo request or a UDP packet.
- During an SNMP test, one probe operation means sending one SNMPv1 packet, one SNMPv2C packet, and one SNMPv3 packet.

NQA client and server

A device with NQA test groups configured is an NQA client, and the NQA client initiates NQA tests. An NQA server makes responses to probe packets destined to the specified destination address and port number.

Figure 39 Relationship between the NQA client and NQA server



Not all test types require the NQA server. Only the TCP, UDP echo, UDP jitter, or voice test requires both the NQA client and server, as shown in [Figure 39](#).

You can create multiple TCP or UDP listening services on the NQA server. Each listens to a specific destination address and port number. Make sure the destination IP address and port number for a listening service on the server are the same as those configured for the test group on the NQA client. Each listening service must be unique on the NQA server.

NQA probe operation procedure

An NQA probe operation involves the following steps:

1. The NQA client constructs probe packets for the specified type of NQA test, and sends them to the peer device.
2. Upon receiving the probe packets, the peer sends back responses with timestamps.
3. The NQA client computes the network performance and service quality parameters, such as the packet loss rate and round-trip time based on the received responses.

NQA configuration task list

| Task | Remarks |
|--|---|
| Configuring the NQA server | Required for TCP, UDP echo, UDP jitter, and voice tests |

To perform NQA tests successfully, perform the following configurations on the NQA client:

1. Enable the NQA client.
2. Create a test group and configure test parameters. The test parameters may vary with test types.
3. Configure a schedule for the NQA test group.

Complete these tasks to configure NQA client:

| Task | Remarks | |
|---|---|----------------------------|
| Enabling the NQA client | Required. | |
| Creating an NQA test group | Required. | |
| Configuring an NQA test group | Configuring ICMP echo tests | Required. |
| | Configuring DHCP tests | Use any of the approaches. |
| | Configuring DNS tests | |

| Task | Remarks |
|---|-----------|
| Configuring FTP tests | |
| Configuring HTTP tests | |
| Configuring UDP jitter tests | |
| Configuring SNMP tests | |
| Configuring TCP tests | |
| Configuring UDP echo tests | |
| Configuring voice tests | |
| Configuring DLSw tests | |
| Configuring the collaboration function | Optional. |
| Configuring threshold monitoring | Optional. |
| Configuring the NQA statistics collection function | Optional. |
| Configuring the history records saving function | Optional. |
| Configuring optional parameters for an NQA test group | Optional. |
| Configuring a schedule for an NQA test group | Required. |

Configuring the NQA server

To perform TCP, UDP echo, UDP jitter, or voice tests, configure the NQA server on the peer device. The NQA server responds to the probe packets sent from the NQA client by listening to the specified destination address and port number.

To configure the NQA server:

| Step | Command | Remarks |
|---|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enable the NQA server. | nqa server enable | Disabled by default. |
| 3. Configure the listening service. | nqa server { tcp-connect udp-echo } ip-address port-number | The destination IP address and port number must be the same as those configured on the NQA client. A listening service must be unique on the NQA server. |
| 4. Configure the ToS value in the packets sent by the TCP or UDP listening service on the NQA server. | nqa server { tcp-connect udp-echo } tos tos | Optional.
By default, the ToS value is 0. |

Enabling the NQA client

Configurations on the NQA client take effect only when the NQA client is enabled.

To enable the NQA client:

| Step | Command | Remarks |
|---------------------------|-------------------------|----------------------------------|
| 1. Enter system view. | system-view | N/A |
| 2. Enable the NQA client. | nqa agent enable | Optional.
Enabled by default. |

Creating an NQA test group

Create an NQA test group before you configure NQA tests.

To create an NQA test group:

| Step | Command | Remarks |
|---|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Create an NQA test group, and enter the NQA test group view. | nqa entry <i>admin-name</i>
<i>operation-tag</i> | In the NQA test group view, you can specify the test type
You can use the nqa entry command to enter the test type view of an NQA test group with test type configured. |

Configuring an NQA test group

Configuring ICMP echo tests

ICMP echo tests of an NQA test group uses ICMP echo response information to test reachability of a destination host. An ICMP echo test has the same function as the **ping** command but provides more output information. In addition, you can specify the next hop for ICMP echo tests. ICMP echo tests are used to locate connectivity problems in a network.

NQA ICMP echo tests are not supported in IPv6 networks. To test the reachability of an IPv6 address, use the **ping ipv6** command. For more information about the command, see *Network Management and Monitoring Command Reference*.

To configure ICMP echo tests:

| Step | Command | Remarks |
|--|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter NQA test group view. | nqa entry <i>admin-name</i>
<i>operation-tag</i> | N/A |
| 3. Configure the test type as ICMP echo, and enter test type view. | type icmp-echo | N/A |
| 4. Configure the destination address of ICMP echo requests. | destination ip <i>ip-address</i> | By default, no destination IP address is configured. |
| 5. Configure the size of the data field in each ICMP echo request. | data-size <i>size</i> | Optional.
100 bytes by default. |

| Step | Command | Remarks |
|---|---|--|
| 6. Configure the string to be filled in the data field of each ICMP echo request. | data-fill <i>string</i> | Optional.
By default, the string is the hexadecimal number 00010203040506070809. |
| 7. Apply ICMP echo tests to the specified VPN. | vpn-instance
<i>vpn-instance-name</i> | Optional.
By default, ICMP echo tests apply to the public network.
Only the HP 5500 EI Switch Series supports VPN instances. |
| 8. Configure the source interface for ICMP echo requests. | source interface
<i>interface-type</i>
<i>interface-number</i> | Optional.
By default, no source interface is configured for probe packets.
The requests take the IP address of the source interface as their source IP address when no source IP address is specified.
The specified source interface must be up; otherwise, no ICMP echo requests can be sent out. |
| 9. Configure the source IP address of ICMP echo requests. | source ip <i>ip-address</i> | Optional.
By default, no source IP address is configured.
If you configure both the source ip command and the source interface command, the source ip command takes effect.
The source IP address must be the IP address of a local interface. The local interface must be up; otherwise, no ICMP echo requests can be sent out. |
| 10. Configure the next hop IP address of ICMP echo requests. | next-hop <i>ip-address</i> | Optional.
By default, no next hop IP address is configured. |
| 11. Configure optional parameters. | See " Configuring optional parameters for an NQA test group " | Optional. |

Configuring DHCP tests

DHCP tests of an NQA test group are used to test if a DHCP server is on the network, and the time for the DHCP server to respond to a client request and assign an IP address to the client.

Before you start DHCP tests, configure the DHCP server. If the NQA (DHCP client) and the DHCP server are not in the same network segment, configure a DHCP relay. For the configuration of DHCP server and DHCP relay, see *Layer 3—IP Services Configuration Guide*.

The interface that performs DHCP tests does not change its IP address. A DHCP test only simulates address allocation in DHCP.

When a DHCP test completes, the NQA client sends a DHCP-RELEASE packet to release the obtained IP address.

To configuring DHCP tests:

| Step | Command | Remarks |
|---|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter NQA test group view. | nqa entry <i>admin-name</i>
<i>operation-tag</i> | N/A |
| 3. Configure the test type as DHCP, and enter test type view. | type dhcp | N/A |
| 4. Specify an interface to perform DHCP tests. | operation interface
<i>interface-type</i>
<i>interface-number</i> | By default, no interface is configured to perform DHCP tests.
The specified interface must be up; otherwise, no probe packets can be sent out. |
| 5. Configure optional parameters. | See " Configuring optional parameters for an NQA test group " | Optional. |

Configuring DNS tests

DNS tests of an NQA test group are used to test whether the NQA client can translate a domain name into an IP address through a DNS server and test the time required for resolution.

Before you start DNS tests, configure the mapping between a domain name and an IP address on a DNS server.

A DNS test simulates the domain name resolution. It does not save the mapping between the domain name and the IP address.

To configure DNS tests:

| Step | Command | Remarks |
|--|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter NQA test group view. | nqa entry <i>admin-name</i> <i>operation-tag</i> | N/A |
| 3. Configure the test type as DNS, and enter test type view. | type dns | N/A |
| 4. Specify the IP address of the DNS server as the destination address of DNS packets. | destination ip <i>ip-address</i> | By default, no destination IP address is configured. |
| 5. Configure the domain name that needs to be translated. | resolve-target <i>domain-name</i> | By default, no domain name is configured. |
| 6. Configure optional parameters. | See " Configuring optional parameters for an NQA test group " | Optional. |

Configuring FTP tests

FTP tests of an NQA test group are used to test the connection between the NQA client and an FTP server and the time required for the FTP client to transfer a file to or download a file from the FTP server.

Before you start FTP tests, configure the FTP server. For example, configure a username and password that are used to log in to the FTP server. For more information about FTP server configuration, see *Fundamentals Configuration Guide*.

Follow these guidelines when you configure FTP tests:

- When you execute the **put** command, the NQA client creates a file named *file-name* of fixed size on the FTP server. The *file-name* argument does not represent any file on the NQA client. When you execute the **get** command, the client does not save the files obtained from the FTP server.
- When you get a file that does not exist on the FTP server, FTP tests fail.
- When you perform an FTP test, use a small file and set a long NQA probe timeout time. A big file or short probe timeout time may result in probe timeout.

To configure FTP tests:

| Step | Command | Remarks |
|--|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter NQA test group view. | nqa entry <i>admin-name</i>
<i>operation-tag</i> | N/A |
| 3. Configure the test type as FTP, and enter test type view. | type ftp | N/A |
| 4. Specify the IP address of the FTP server as the destination address of FTP request packets. | destination ip <i>ip-address</i> | By default, no destination IP address is configured. |
| 5. Configure the source IP address of FTP request packets. | source ip <i>ip-address</i> | By default, no source IP address is specified.
The source IP address must be the IP address of a local interface. The local interface must be up; otherwise, no FTP requests can be sent out. |
| 6. Configure the operation type. | operation { get put } | Optional.
By default, the operation type for the FTP is get , which means obtaining files from the FTP server. |
| 7. Configure a login username. | username <i>name</i> | By default, no login username is configured. |
| 8. Configure a login password. | password [cipher simple] <i>password</i> | By default, no login password is configured. |
| 9. Specify a file to be transferred between the FTP server and the FTP client. | filename <i>file-name</i> | By default, no file is specified. |
| 10. Set the data transmission mode for FTP tests. | mode { active passive } | Optional.
active by default. |

| Step | Command | Remarks |
|------------------------------------|---|-----------|
| 11. Configure optional parameters. | See " Configuring optional parameters for an NQA test group " | Optional. |

Configuring HTTP tests

HTTP tests of an NQA test group are used to test the connection between the NQA client and an HTTP server, and the time required to obtain data from the HTTP server. HTTP tests enable you to detect the connectivity and performance of the HTTP server. The TCP port must be port 80 on the HTTP server for NQA HTTP tests.

Before you start HTTP tests, configure the HTTP server.

To configure HTTP tests:

| Step | Command | Remarks |
|--|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter NQA test group view. | nqa entry <i>admin-name</i>
<i>operation-tag</i> | N/A |
| 3. Configure the test type as HTTP, and enter test type view. | type http | N/A |
| 4. Configure the IP address of the HTTP server as the destination address of HTTP request packets. | destination ip <i>ip-address</i> | By default, no destination IP address is configured. |
| 5. Configure the source IP address of request packets. | source ip <i>ip-address</i> | Optional.
By default, no source IP address is specified.
The source IP address must be the IP address of a local interface. The local interface must be up; otherwise, no probe packets can be sent out. |
| 6. Configure the operation type. | operation { get post } | Optional.
By default, the operation type for the HTTP is get , which means obtaining data from the HTTP server. |
| 7. Configure the website that an HTTP test visits. | url <i>url</i> | N/A |
| 8. Configure the HTTP version used in HTTP tests. | http-version v1.0 | Optional.
By default, HTTP 1.0 is used. |
| 9. Configure optional parameters. | See " Configuring optional parameters for an NQA test group " | Optional. |

Configuring UDP jitter tests

! **IMPORTANT:**

Do not perform NQA UDP jitter tests on known ports, ports from 1 to 1023. Otherwise, UDP jitter tests might fail or the corresponding services of this port might be unavailable.

Real-time services such as voice and video have high requirements on delay jitters. UDP jitter tests of an NQA test group obtain uni/bi-directional delay jitters. The test results help you verify whether a network can carry real-time services.

A UDP jitter test performs the following procedure:

1. The source sends packets to the destination port at regular intervals.
2. The destination affixes a time stamp to each packet that it receives, and then sends the packet back to the source.
3. Upon receiving the response, the source calculates the delay jitter, which reflects network performance. Delay refers to the amount of time it takes a packet to be transmitted from source to destination or from destination to source. Delay jitter is the delay variation over time.

Configuration prerequisites

UDP jitter tests require cooperation between the NQA server and the NQA client. Before you start UDP jitter tests, configure UDP listening services on the NQA server. For more information about UDP listening service configuration, see "[Configuring the NQA server.](#)"

Configuration procedure

To configure UDP jitter tests:

| Step | Command | Remarks |
|--|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter NQA test group view. | nqa entry <i>admin-name</i>
<i>operation-tag</i> | N/A |
| 3. Configure the test type as UDP jitter, and enter test type view. | type udp-jitter | N/A |
| 4. Configure the destination address of UDP packets. | destination ip <i>ip-address</i> | By default, no destination IP address is configured.
The destination IP address must be the same as that of the listening service on the NQA server. |
| 5. Configure the destination port of UDP packets. | destination port
<i>port-number</i> | By default, no destination port number is configured.
The destination port must be the same as that of the listening service on the NQA server. |
| 6. Specify the source port number of UDP packets. | source port <i>port-number</i> | Optional.
By default, no source port number is specified. |
| 7. Configure the size of the data field in each UDP packet. | data-size <i>size</i> | Optional.
100 bytes by default. |
| 8. Configure the string to be filled in the data field of each probe packet. | data-fill <i>string</i> | Optional.
By default, the string is the hexadecimal number 00010203040506070809. |

| Step | Command | Remarks |
|---|---|--|
| 9. Configure the number of probe packets to be sent during each UDP jitter probe operation. | probe packet-number
<i>packet-number</i> | Optional.
10 by default. |
| 10. Configure the interval for sending probe packets during each UDP jitter probe operation. | probe packet-interval
<i>packet-interval</i> | Optional.
20 milliseconds by default. |
| 11. Configure the interval the NQA client must wait for a response from the server before it regards the response is timed out. | probe packet-timeout
<i>packet-timeout</i> | Optional.
3000 milliseconds by default. |
| 12. Configure the source IP address for UDP jitter packets. | source ip <i>ip-address</i> | Optional.
By default, no source IP address is specified.
The source IP address must be the IP address of a local interface. The local interface must be up; otherwise, no probe packets can be sent out. |
| 13. Configure optional parameters. | See " Configuring optional parameters for an NQA test group " | Optional. |

NOTE:

The **probe count** command specifies the number of probe operations during one UDP jitter test. The **probe packet-number** command specifies the number of probe packets sent in each UDP jitter probe operation.

Configuring SNMP tests

SNMP tests of an NQA test group are used to test the time the NQA client takes to send an SNMP packet to the SNMP agent and receive a response.

Before you start SNMP tests, enable the SNMP agent function on the device that serves as an SNMP agent. For more information about SNMP agent configuration, see "Configuring SNMP."

To configure SNMP tests:

| Step | Command | Remarks |
|---|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter NQA test group view. | nqa entry <i>admin-name</i>
<i>operation-tag</i> | N/A |
| 3. Configure the test type as SNMP, and enter test type view. | type snmp | N/A |
| 4. Configure the destination address of SNMP packets. | destination ip <i>ip-address</i> | By default, no destination IP address is configured. |

| Step | Command | Remarks |
|---|---|--|
| 5. Specify the source port of SNMP packets. | source port <i>port-number</i> | Optional.
By default, no source port number is specified. |
| 6. Configure the source IP address of SNMP packets. | source ip <i>ip-address</i> | Optional.
By default, no source IP address is specified.
The source IP address must be the IP address of a local interface. The local interface must be up; otherwise, no probe packets can be sent out. |
| 7. Configure optional parameters. | See " Configuring optional parameters for an NQA test group " | Optional. |

Configuring TCP tests

TCP tests of an NQA test group are used to test the TCP connection between the NQA client and a port on the NQA server and the time for setting up a connection. The test result helps you evaluate the availability and performance of the services provided by the port on the server.

TCP tests require cooperation between the NQA server and the NQA client. Before you start TCP tests, configure a TCP listening service on the NQA server. For more information about the TCP listening service configuration, see "[Configuring the NQA server.](#)"

To configure TCP tests:

| Step | Command | Remarks |
|--|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter NQA test group view. | nqa entry <i>admin-name</i>
<i>operation-tag</i> | N/A |
| 3. Configure the test type as TCP, and enter test type view. | type tcp | N/A |
| 4. Configure the destination address of TCP probe packets. | destination ip <i>ip-address</i> | By default, no destination IP address is configured.
The destination address must be the same as the IP address of the listening service configured on the NQA server. |
| 5. Configure the destination port of TCP probe packets. | destination port
<i>port-number</i> | By default, no destination port number is configured.
The destination port number must be the same as that of the listening service on the NQA server. |
| 6. Configure the source IP address of TCP probe packets. | source ip <i>ip-address</i> | Optional.
By default, no source IP address is specified.
The source IP address must be the IP address of a local interface. The local interface must be up; otherwise, no probe packets can be sent out. |
| 7. Configure optional parameters. | See " Configuring optional parameters for an NQA test group " | Optional. |

Configuring UDP echo tests

UDP echo tests of an NQA test group are used to test the connectivity and round-trip time of a UDP packet from the client to the specified UDP port on the NQA server.

UDP echo tests require cooperation between the NQA server and the NQA client. Before you start UDP echo tests, configure a UDP listening service on the NQA server. For more information about the UDP listening service configuration, see "[Configuring the NQA server.](#)"

To configure UDP echo tests:

| Step | Command | Remarks |
|--|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter NQA test group view. | nqa entry <i>admin-name</i>
<i>operation-tag</i> | N/A |
| 3. Configure the test type as UDP echo, and enter test type view. | type udp-echo | N/A |
| 4. Configure the destination address of UDP packets. | destination ip <i>ip-address</i> | By default, no destination IP address is configured.
The destination address must be the same as the IP address of the listening service configured on the NQA server. |
| 5. Configure the destination port of UDP packets. | destination port <i>port-number</i> | By default, no destination port number is configured.
The destination port number must be the same as that of the listening service on the NQA server. |
| 6. Configure the size of the data field in each UDP packet. | data-size <i>size</i> | Optional.
100 bytes by default. |
| 7. Configure the string to be filled in the data field of each UDP packet. | data-fill <i>string</i> | Optional.
By default, the string is the hexadecimal number 00010203040506070809. |
| 8. Specify the source port of UDP packets. | source port <i>port-number</i> | Optional.
By default, no source port number is specified. |
| 9. Configure the source IP address of UDP packets. | source ip <i>ip-address</i> | Optional.
By default, no source IP address is specified.
The source IP address must be that of an interface on the device and the interface must be up; otherwise, no probe packets can be sent out. |
| 10. Configure optional parameters. | See " Configuring optional parameters for an NQA test group " | Optional. |

Configuring voice tests

ⓘ **IMPORTANT:**

Do not perform voice tests on known ports, ports from 1 to 1023. Otherwise, the NQA test might fail or the corresponding services of these ports might be unavailable.

Voice tests of an NQA test group are used to test voice over IP (VoIP) network status, and collect VoIP network parameters so that users can adjust the network.

A voice test performs the following procedure:

1. The source (NQA client) sends voice packets of G.711 A-law, G.711 μ -law or G.729 A-law codec type at regular intervals to the destination (NQA server).
2. The destination affixes a time stamp to each voice packet that it receives and then sends it back to the source.
3. Upon receiving the packet, the source calculates results, such as the delay jitter and one-way delay based on the packet time stamps. The statistics reflect network performance.

Voice test result also includes the following parameters that reflect VoIP network performance:

- Calculated Planning Impairment Factor (ICPIF)—Measures impairment to voice quality in a VoIP network. It is decided by packet loss and delay. A higher value represents a lower service quality.
- Mean Opinion Scores (MOS)—A MOS value can be evaluated by using the ICPIF value, in the range of 1 to 5. A higher value represents a higher quality of a VoIP network.

The evaluation of voice quality depends on users' tolerance for voice quality, which should be taken into consideration. For users with higher tolerance for voice quality, use the **advantage-factor** command to configure the advantage factor. When the system calculates the ICPIF value, this advantage factor is subtracted to modify ICPIF and MOS values, so objective and subjective factors are both considered when you evaluate voice quality.

Configuration prerequisites

- Voice tests require cooperation between the NQA server and the NQA client. Before you start voice tests, configure a UDP listening service on the NQA server. For more information about UDP listening service configuration, see "[Configuring the NQA server.](#)"
- Only one probe operation is performed in one voice test.

Configuration procedure

To configure voice tests:

| Step | Command | Remarks |
|--|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter NQA test group view. | nqa entry <i>admin-name</i>
<i>operation-tag</i> | N/A |
| 3. Configure the test type as voice, and enter test type view. | type voice | N/A |
| 4. Configure the destination address of voice probe packets. | destination ip <i>ip-address</i> | By default, no destination IP address is configured for a test operation.
The destination IP address must be the same as that of the listening service on the NQA server. |

| Step | Command | Remarks |
|--|---|--|
| 5. Configure the destination port of voice probe packets. | destination port <i>port-number</i> | By default, no destination port number is configured.
The destination port must be the same as that of the listening service on the NQA server. |
| 6. Configure the codec type. | codec-type { g711a g711u g729a } | Optional.
By default, the codec type is G.711 A-law. |
| 7. Configure the advantage factor for calculating MOS and ICPIF values. | advantage-factor <i>factor</i> | Optional.
By default, the advantage factor is 0. |
| 8. Specify the source IP address of probe packets. | source ip <i>ip-address</i> | Optional.
By default, no source IP address is specified.
The source IP address must be the IP address of a local interface. The local interface must be up; otherwise, no probe packets can be sent out. |
| 9. Specify the source port number of probe packets. | source port <i>port-number</i> | Optional.
By default, no source port number is specified. |
| 10. Configure the size of the data field in each probe packet. | data-size <i>size</i> | Optional.
By default, the probe packet size depends on the codec type. The default packet size is 172 bytes for G.711A-law and G.711 μ -law codec type, and 32 bytes for G.729 A-law codec type. |
| 11. Configure the string to be filled in the data field of each probe packet. | data-fill <i>string</i> | Optional.
By default, the string is the hexadecimal number 00010203040506070809. |
| 12. Configure the number of probe packets to be sent during each voice probe operation. | probe packet-number <i>packet-number</i> | Optional.
1000 by default. |
| 13. Configure the interval for sending probe packets during each voice probe operation. | probe packet-interval <i>packet-interval</i> | Optional.
20 milliseconds by default. |
| 14. Configure the interval the NQA client must wait for a response from the server before it regards the response times out. | probe packet-timeout <i>packet-timeout</i> | Optional.
5000 milliseconds by default. |
| 15. Configure optional parameters. | See " Configuring optional parameters for an NQA test group " | Optional. |

Configuring DLSw tests

DLSw tests of an NQA test group are used to test the response time of a DLSw device.

Before you start DLSw tests, enable the DLSw function on the peer device.

To configure DLSw tests:

| Step | Command | Remarks |
|---|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter NQA test group view. | nqa entry <i>admin-name</i>
<i>operation-tag</i> | N/A |
| 3. Configure the test type as DLSw, and enter test type view. | type dlsw | N/A |
| 4. Configure the destination address of probe packets. | destination ip <i>ip-address</i> | By default, no destination IP address is configured. |
| 5. Configure the source IP address of probe packets. | source ip <i>ip-address</i> | Optional.
By default, no source IP address is specified.
The source IP address must be the IP address of a local interface. The local interface must be up; otherwise, no probe packets can be sent out. |
| 6. Configure optional parameters. | See " Configuring optional parameters for an NQA test group " | Optional. |

Configuring the collaboration function

Collaboration is implemented by establishing reaction entries to monitor the detection results of a test group. If the number of consecutive probe failures reaches the threshold, the configured action is triggered.

To configure the collaboration function:

| Step | Command | Remarks |
|--|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter NQA test group view. | nqa entry <i>admin-name</i>
<i>operation-tag</i> | N/A |
| 3. Enter test type view of the test group. | type { dhcp dlsw dns ftp http icmp-echo snmp tcp udp-echo } | The collaboration function is not supported in UDP jitter and voice tests. |
| 4. Configure a reaction entry. | reaction <i>item-number</i>
checked-element probe-fail
threshold-type consecutive
<i>consecutive-occurrences</i>
action-type trigger-only | Not created by default.
You cannot modify the content of an existing reaction entry. |
| 5. Exit to system view. | quit | N/A |
| 6. Configure a track entry and associate it with the reaction entry of the NQA test group. | track <i>entry-number</i> nqa entry
<i>admin-name operation-tag</i>
reaction <i>item-number</i> | Not created by default. |

Configuring threshold monitoring

Configuration prerequisites

Before you configure threshold monitoring, complete the following tasks:

- Configure the destination address of the trap message by using the **snmp-agent target-host** command. For more information about the **snmp-agent target-host** command, see *Network Management and Monitoring Command Reference*.
- Create an NQA test group and configure the related parameters.

Configuration guidelines

Follow these guidelines when you configure threshold monitoring:

- NQA DNS tests do not support the action of sending trap messages. The action to be triggered in DNS tests can only be the default one, **none**.
- Only the **test-complete** keyword is supported for the **reaction trap** command in a voice test.

Configuration procedure

To configure threshold monitoring:

| Step | Command | Remarks |
|--|---|---------|
| 1. Enter system view. | system-view | N/A |
| 2. Enter NQA test group view. | nqa entry <i>admin-name operation-tag</i> | N/A |
| 3. Enter test type view of the test group. | type { dhcp dls w dns ftp http icmp-echo snmp tcp udp-echo udp-jitter voice } | N/A |

| Step | Command | Remarks |
|------------------------------------|--|---|
| 4. Configure threshold monitoring. | <ul style="list-style-type: none"> • Enable sending traps to the network management server under specified conditions:
reaction trap { probe-failure <i>consecutive-probe-failures</i> test-complete test-failure <i>cumulate-probe-failures</i> } • Configure a reaction entry for monitoring the probe duration of a test (not supported in UDP jitter and voice tests):
reaction <i>item-number</i> checked-element probe-duration threshold-type { accumulate <i>accumulate-occurrences</i> average consecutive <i>consecutive-occurrences</i> } threshold-value <i>upper-threshold lower-threshold</i> [action-type { none trap-only }] • Configure a reaction entry for monitoring the probe failure times (not supported in UDP jitter and voice tests):
reaction <i>item-number</i> checked-element probe-fail threshold-type { accumulate <i>accumulate-occurrences</i> consecutive <i>consecutive-occurrences</i> } [action-type { none trap-only }] • Configure a reaction entry for monitoring packet round-trip time (only supported in UDP jitter and voice tests):
reaction <i>item-number</i> checked-element rtt threshold-type { accumulate <i>accumulate-occurrences</i> average } threshold-value <i>upper-threshold lower-threshold</i> [action-type { none trap-only }] • Configure a reaction entry for monitoring the packet loss in each test (only supported in UDP jitter and voice tests):
reaction <i>item-number</i> checked-element packet-loss threshold-type accumulate <i>accumulate-occurrences</i> [action-type { none trap-only }] • Configure a reaction entry for monitoring one-way delay jitter (only supported in UDP jitter and voice tests):
reaction <i>item-number</i> checked-element { jitter-ds jitter-sd } threshold-type { accumulate <i>accumulate-occurrences</i> average } threshold-value <i>upper-threshold lower-threshold</i> [action-type { none trap-only }] • Configure a reaction entry for monitoring the one-way delay (only supported in UDP jitter and voice tests):
reaction <i>item-number</i> checked-element { owd-ds owd-sd } threshold-value <i>upper-threshold lower-threshold</i> • Configure a reaction entry for monitoring the ICPIF value (only supported in voice tests):
reaction <i>item-number</i> checked-element icpif threshold-value <i>upper-threshold lower-threshold</i> [action-type { none trap-only }] • Configure a reaction entry for monitoring the MOS value (only supported in voice tests):
reaction <i>item-number</i> checked-element mos threshold-value <i>upper-threshold lower-threshold</i> [action-type { none trap-only }] | <p>Configure to send traps.</p> <p>No traps are sent to the network management server by default.</p> |

Configuring the NQA statistics collection function

NQA groups tests completed in a time period for a test group, and calculates the test result statistics. The statistics form a statistics group. To view information about the statistics groups, use the **display nqa statistics** command. To set the interval for collecting statistics, use the **statistics interval** command.

When the number of statistics groups kept reaches the upper limit and a new statistics group is to be saved, the oldest statistics group is deleted. To set the maximum number of statistics groups that can be kept, use the **statistics max-group** command.

A statistics group is formed after the last test is completed within the specified interval. Statistics groups have an aging mechanism. A statistics group is deleted when its hold time expires. To set the hold time of statistics groups for a test group, use the **statistics hold-time** command.

Follow these guidelines when you configure the NQA statistics collection function:

- The NQA statistics collection function is not supported in DHCP tests.
- If you use the **frequency** command to set the frequency between two consecutive tests to 0, only one test is performed, and no statistics group information is collected.

To configure the NQA statistics collection function:

| Step | Command | Remarks |
|--|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter NQA test group view. | nqa entry <i>admin-name operation-tag</i> | N/A |
| 3. Enter test type view of the test group. | type { <i>dls</i> w <i>dns</i> <i>ftp</i> <i>http</i> <i>icmp-echo</i> <i>snmp</i> <i>tcp</i> <i>udp-echo</i> <i>udp-jitter</i> <i>voice</i> } | N/A |
| 4. Configure the interval for collecting the statistics of test results. | statistics interval <i>interval</i> | Optional.
60 minutes by default. |
| 5. Configure the maximum number of statistics groups that can be kept. | statistics max-group <i>number</i> | Optional.
2 by default.
To disable collecting NQA statistics, set the maximum number to 0. |
| 6. Configure the hold time of statistics groups. | statistics hold-time <i>hold-time</i> | Optional.
120 minutes by default. |

Configuring the history records saving function

The history records saving function enables the system to save the history records of NQA tests. To view the history records of a test group, use the **display nqa history** command.

The configuration task also allows you to configure the following elements:

- **Lifetime of the history records**—The records are removed when the lifetime is reached.
- **The maximum number of history records that can be saved in a test group**—If the number of history records in a test group exceeds the maximum number, the earliest history records are removed.

To configure the history records saving function of an NQA test group:

| Step | Command | Remarks |
|--|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter NQA test group view. | nqa entry <i>admin-name</i>
<i>operation-tag</i> | N/A |
| 3. Enter NQA test type view. | type { dhcp dls w dns ftp
 http icmp-echo snmp
tcp udp-echo udp-jitter
voice } | N/A |
| 4. Enable the saving of the history records of the NQA test group. | history-record enable | By default, history records of the NQA test group are not saved. |
| 5. Set the lifetime of the history records in an NQA test group. | history-record keep-time
<i>keep-time</i> | Optional.
By default, the history records in the NQA test group are kept for 120 minutes. |
| 6. Configure the maximum number of history records that can be saved for a test group. | history-record number
<i>number</i> | Optional.
By default, the maximum number of records that can be saved for a test group is 50. |

Configuring optional parameters for an NQA test group

Optional parameters for an NQA test group are valid only for tests in this test group.

Unless otherwise specified, the following optional parameters are applicable to all test types.

To configure optional parameters for an NQA test group:

| Step | Command | Remarks |
|--|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter NQA test group view. | nqa entry <i>admin-name</i>
<i>operation-tag</i> | N/A |
| 3. Enter test type view of a test group. | type { dhcp dls w dns
 ftp http icmp-echo
 snmp tcp udp-echo
 udp-jitter voice } | N/A |
| 4. Configure the description for a test group. | description <i>text</i> | Optional.
By default, no description is available for a test group. |

| Step | Command | Remarks |
|--|--|---|
| 5. Configure the interval between two consecutive tests for a test group. | frequency <i>interval</i> | Optional.
By default, the interval between two consecutive tests for a test group is 0 milliseconds. Only one test is performed.
If the last test is not completed when the interval specified by the frequency command is reached, a new test does not start. |
| 6. Configure the number of probe operations to be performed in one test. | probe count <i>times</i> | Optional.
By default, one probe operation is performed in one test.
Not available for voice tests, Only one probe operation can be performed in one voice test. |
| 7. Configure the NQA probe timeout time. | probe timeout <i>timeout</i> | Optional.
By default, the timeout time is 3000 milliseconds.
Not available for UDP jitter tests. |
| 8. Configure the maximum number of hops a probe packet traverses in the network. | ttl <i>value</i> | Optional.
20 by default.
Not available for DHCP tests. |
| 9. Configure the ToS field in an IP packet header in an NQA probe packet. | tos <i>value</i> | Optional.
0 by default. |
| 10. Enable the routing table bypass function. | route-option
bypass-route | Optional.
Disabled by default.
Not available for DHCP tests. |

Configuring a schedule for an NQA test group

You can configure a schedule for an NQA test group by setting the start time and test duration for a test group.

A test group performs tests between the scheduled start time and the end time (the start time plus test duration). If the scheduled start time is ahead of the system time, the test group starts testing immediately. If both the scheduled start and end time are behind the system time, no test will start. To view the current system time, use the **display clock** command.

Configuration prerequisites

Before you configure a schedule for an NQA test group, complete the following tasks:

- Configure test parameters required for the test type.
- Configure the NQA server for tests that require cooperation with the NQA server.

Configuration guidelines

Follow these guidelines when you schedule an NQA test group:

- After an NQA test group is scheduled, you cannot enter the test group view or test type view.
- System adjustment does not affect started or completed test groups. It only affects test groups that have not started.

Configuration procedure

To schedule an NQA test group:

| Step | Command | Remarks |
|--|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Configure a schedule for an NQA test group. | nqa schedule <i>admin-name operation-tag start-time</i> { <i>hh:mm:ss</i> [<i>yyyy/mm/dd</i>] now } lifetime { <i>lifetime</i> forever } | now specifies the test group starts testing immediately.
forever specifies that the tests do not stop unless you use the undo nqa schedule command. |
| 3. Configure the maximum number of tests that the NQA client can simultaneously perform. | nqa agent max-concurrent <i>number</i> | Optional.
By default, the maximum number of tests that the NQA client can simultaneously perform is 2. |

Displaying and maintaining NQA

| Task | Command | Remarks |
|--|--|-----------------------|
| Display history records of NQA test groups. | display nqa history [<i>admin-name operation-tag</i>] [{ begin exclude include } <i>regular-expression</i>] | |
| Display the current monitoring results of reaction entries. | display nqa reaction counters [<i>admin-name operation-tag</i> [<i>item-number</i>]] [{ begin exclude include } <i>regular-expression</i>] | |
| Display the results of the last NQA test. | display nqa result [<i>admin-name operation-tag</i>] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display statistics of test results for the specified or all test groups. | display nqa statistics [<i>admin-name operation-tag</i>] [{ begin exclude include } <i>regular-expression</i>] | |
| Display NQA server status. | display nqa server status [{ begin exclude include } <i>regular-expression</i>] | |

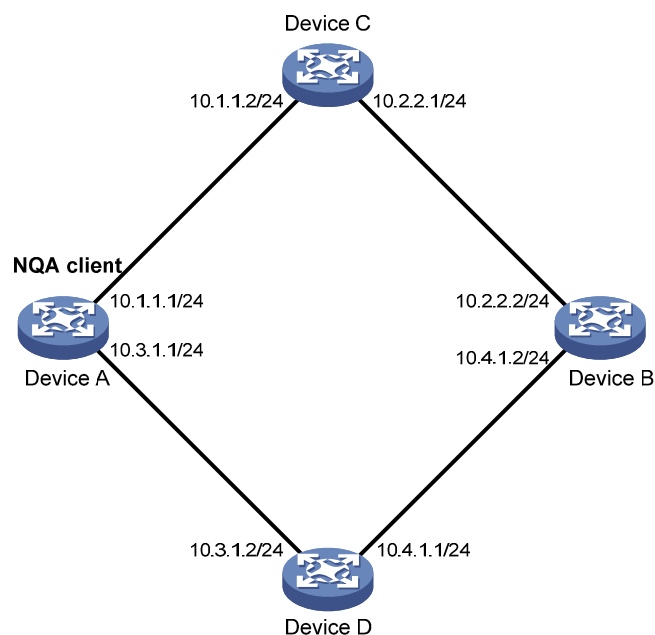
NQA configuration examples

ICMP echo test configuration example

Network requirements

As shown in [Figure 40](#), configure NQA ICMP echo tests to test whether the NQA client (Device A) can send packets through a specific next hop to the specified destination (Device B) and test the round-trip time of the packets.

Figure 40 Network diagram



Configuration procedure

Before you make the configuration, make sure the devices can reach each other.

Create an ICMP echo test group, and specify 10.2.2.2 as the destination IP address for ICMP echo requests to be sent.

```
<DeviceA> system-view
[DeviceA] nqa entry admin test
[DeviceA-nqa-admin-test] type icmp-echo
[DeviceA-nqa-admin-test-icmp-echo] destination ip 10.2.2.2
```

Configure 10.1.1.2 as the next hop IP address for ICMP echo requests. The ICMP echo requests are sent to Device C to Device B (the destination).

```
[DeviceA-nqa-admin-test-icmp-echo] next-hop 10.1.1.2
```

Configure the device to perform 10 probe operations per test, perform tests at an interval of 5000 milliseconds. Set the NQA probe timeout time as 500 milliseconds.

```
[DeviceA-nqa-admin-test-icmp-echo] probe count 10
[DeviceA-nqa-admin-test-icmp-echo] probe timeout 500
[DeviceA-nqa-admin-test-icmp-echo] frequency 5000
```

Enable the saving of history records and configure the maximum number of history records that can be saved for a test group.

```
[DeviceA-nqa-admin-test-icmp-echo] history-record enable
[DeviceA-nqa-admin-test-icmp-echo] history-record number 10
[DeviceA-nqa-admin-test-icmp-echo] quit
```

Start ICMP echo tests.

```
[DeviceA] nqa schedule admin test start-time now lifetime forever
```

Stop the ICMP echo tests after a period of time.

```
[DeviceA] undo nqa schedule admin test
```

Display the results of the last ICMP echo test.

```
[DeviceA] display nqa result admin test
NQA entry (admin admin, tag test) test results:
  Destination IP address: 10.2.2.2
  Send operation times: 10          Receive response times: 10
  Min/Max/Average round trip time: 2/5/3
  Square-Sum of round trip time: 96
  Last succeeded probe time: 2011-01-23 15:00:01.2
Extended results:
  Packet loss in test: 0%
  Failures due to timeout: 0
  Failures due to disconnect: 0
  Failures due to no connection: 0
  Failures due to sequence error: 0
  Failures due to internal error: 0
  Failures due to other errors: 0
  Packet(s) arrived late: 0
```

Display the history of ICMP echo tests.

```
[DeviceA] display nqa history admin test
NQA entry (admin admin, tag test) history record(s):
```

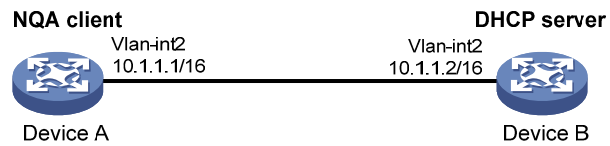
| Index | Response | Status | Time |
|-------|----------|-----------|-----------------------|
| 370 | 3 | Succeeded | 2011-01-23 15:00:01.2 |
| 369 | 3 | Succeeded | 2011-01-23 15:00:01.2 |
| 368 | 3 | Succeeded | 2011-01-23 15:00:01.2 |
| 367 | 5 | Succeeded | 2011-01-23 15:00:01.2 |
| 366 | 3 | Succeeded | 2011-01-23 15:00:01.2 |
| 365 | 3 | Succeeded | 2011-01-23 15:00:01.2 |
| 364 | 3 | Succeeded | 2011-01-23 15:00:01.1 |
| 363 | 2 | Succeeded | 2011-01-23 15:00:01.1 |
| 362 | 3 | Succeeded | 2011-01-23 15:00:01.1 |
| 361 | 2 | Succeeded | 2011-01-23 15:00:01.1 |

DHCP test configuration example

Network requirements

As shown in [Figure 41](#), configure NQA DHCP tests to test the time required for Device A to obtain an IP address from the DHCP server (Device B).

Figure 41 Network diagram



Configuration procedure

Create a DHCP test group, and specify interface VLAN-interface 2 to perform NQA DHCP tests.

```
<DeviceA> system-view
[DeviceA] nqa entry admin test
[DeviceA-nqa-admin-test] type dhcp
[DeviceA-nqa-admin-test-dhcp] operation interface vlan-interface 2
```

Enable the saving of history records.

```
[DeviceA-nqa-admin-test-dhcp] history-record enable
[DeviceA-nqa-admin-test-dhcp] quit
```

Start DHCP tests.

```
[DeviceA] nqa schedule admin test start-time now lifetime forever
```

Stop DHCP tests after a period of time.

```
[DeviceA] undo nqa schedule admin test
```

Display the result of the last DHCP test.

```
[DeviceA] display nqa result admin test
NQA entry (admin admin, tag test) test results:
  Send operation times: 1          Receive response times: 1
  Min/Max/Average round trip time: 624/624/624
  Square-Sum of round trip time: 389376
  Last succeeded probe time: 2011-01-22 09:56:03.2
Extended results:
  Packet loss in test: 0%
  Failures due to timeout: 0
  Failures due to disconnect: 0
  Failures due to no connection: 0
  Failures due to sequence error: 0
  Failures due to internal error: 0
  Failures due to other errors: 0
  Packet(s) arrived late: 0
```

Display the history of DHCP tests.

```
[DeviceA] display nqa history admin test
NQA entry (admin admin, tag test) history record(s):
  Index      Response      Status          Time
  ---      -
  1          624           Succeeded       2011-01-22 09:56:03.2
```

DNS test configuration example

Network requirements

As shown in [Figure 42](#), configure NQA DNS tests to test whether Device A can translate the domain name **host.com** into an IP address through the DNS server and test the time required for resolution.

Figure 42 Network diagram



Configuration procedure

Before you make the configuration, make sure the devices can reach each other.

Create a DNS test group.

```
<DeviceA> system-view
[DeviceA] nqa entry admin test
[DeviceA-nqa-admin-test] type dns
```

Specify the IP address of the DNS server 10.2.2.2 as the destination address for DNS tests, and specify the domain name that needs to be translated as **host.com**.

```
[DeviceA-nqa-admin-test-dns] destination ip 10.2.2.2
[DeviceA-nqa-admin-test-dns] resolve-target host.com
```

Enable the saving of history records.

```
[DeviceA-nqa-admin-test-dns] history-record enable
[DeviceA-nqa-admin-test-dns] quit
```

Start DNS tests.

```
[DeviceA] nqa schedule admin test start-time now lifetime forever
```

Stop the DNS tests after a period of time.

```
[DeviceA] undo nqa schedule admin test
```

Display the results of the last DNS test.

```
[DeviceA] display nqa result admin test
NQA entry (admin admin, tag test) test results:
  Destination IP address: 10.2.2.2
  Send operation times: 1                Receive response times: 1
  Min/Max/Average round trip time: 62/62/62
  Square-Sum of round trip time: 3844
  Last succeeded probe time: 2011-01-10 10:49:37.3
Extended results:
  Packet loss in test: 0%
  Failures due to timeout: 0
  Failures due to disconnect: 0
  Failures due to no connection: 0
  Failures due to sequence error: 0
  Failures due to internal error: 0
  Failures due to other errors: 0
```

```

Packet(s) arrived late: 0
# Display the history of DNS tests.
[DeviceA] display nqa history admin test
NQA entry (admin admin, tag test) history record(s):
  Index      Response      Status      Time
  1          62            Succeeded   2011-01-10 10:49:37.3

```

FTP test configuration example

Network requirements

As shown in [Figure 43](#), configure NQA FTP tests to test the connection with a specific FTP server and the time required for Device A to upload a file to the FTP server. The login username is **admin**, the login password is **systemtest**, and the file to be transferred to the FTP server is **config.txt**.

Figure 43 Network diagram



Configuration procedure

Before you make the configuration, make sure the devices can reach each other.

Create an FTP test group.

```

<DeviceA> system-view
[DeviceA] nqa entry admin test
[DeviceA-nqa-admin-test] type ftp

```

Specify the IP address of the FTP server 10.2.2.2 as the destination IP address for FTP tests.

```

[DeviceA-nqa-admin-test-ftp] destination ip 10.2.2.2

```

Specify 10.1.1.1 as the source IP address for probe packets.

```

[DeviceA-nqa-admin-test-ftp] source ip 10.1.1.1

```

Set the FTP username to **admin**, and password to **systemtest**.

```

[DeviceA-nqa-admin-test-ftp] username admin
[DeviceA-nqa-admin-test-ftp] password systemtest

```

Configure the device to upload file **config.txt** to the FTP server for each probe operation.

```

[DeviceA-nqa-admin-test-ftp] operation put
[DeviceA-nqa-admin-test-ftp] filename config.txt

```

Enable the saving of history records.

```

[DeviceA-nqa-admin-test-ftp] history-record enable
[DeviceA-nqa-admin-test-ftp] quit

```

Start FTP tests.

```

[DeviceA] nqa schedule admin test start-time now lifetime forever

```

Stop the FTP tests after a period of time.

```

[DeviceA] undo nqa schedule admin test

```

Display the results of the last FTP test.

```
[DeviceA] display nqa result admin test
NQA entry (admin admin, tag test) test results:
  Destination IP address: 10.2.2.2
  Send operation times: 1          Receive response times: 1
  Min/Max/Average round trip time: 173/173/173
  Square-Sum of round trip time: 29929
  Last succeeded probe time: 2011-01-22 10:07:28.6
Extended results:
  Packet loss in test: 0%
  Failures due to timeout: 0
  Failures due to disconnect: 0
  Failures due to no connection: 0
  Failures due to sequence error: 0
  Failures due to internal error: 0
  Failures due to other errors: 0
  Packet(s) arrived late: 0
```

Display the history of FTP tests.

```
[DeviceA] display nqa history admin test
NQA entry (admin admin, tag test) history record(s):
  Index      Response      Status          Time
  ---      -
  1         173          Succeeded      2011-01-22 10:07:28.6
```

HTTP test configuration example

Network requirements

As shown in [Figure 44](#), configure NQA HTTP tests to test the connection with a specific HTTP server and the time required to obtain data from the HTTP server.

Figure 44 Network diagram



Configuration procedure

Before you make the configuration, make sure the devices can reach each other.

Create an HTTP test group.

```
<DeviceA> system-view
[DeviceA] nqa entry admin test
[DeviceA-nqa-admin-test] type http
```

Specify the IP address of the HTTP server 10.2.2.2 as the destination IP address for HTTP tests.

```
[DeviceA-nqa-admin-test-http] destination ip 10.2.2.2
```

Configure the device to get data from the HTTP server for each HTTP probe operation. (get is the default HTTP operation type, and this step is optional.)

```
[DeviceA-nqa-admin-test-http] operation get
```

Configure HTTP tests to visit website **/index.htm**.


```

[DeviceA-nqa-admin-test-http] url /index.htm

# Configure the HTTP version 1.0 to be used in HTTP tests. (Version 1.0 is the default version, and this step
is optional.)
[DeviceA-nqa-admin-test-http] http-version v1.0

# Enable the saving of history records.
[DeviceA-nqa-admin-test-http] history-record enable
[DeviceA-nqa-admin-test-http] quit

# Start HTTP tests.
[DeviceA] nqa schedule admin test start-time now lifetime forever

# Stop HTTP tests after a period of time.
[DeviceA] undo nqa schedule admin test

# Display results of the last HTTP test.
[DeviceA] display nqa result admin test
  NQA entry (admin admin, tag test) test results:
    Destination IP address: 10.2.2.2
      Send operation times: 1          Receive response times: 1
      Min/Max/Average round trip time: 64/64/64
      Square-Sum of round trip time: 4096
      Last succeeded probe time: 2011-01-22 10:12:47.9
    Extended results:
      Packet loss in test: 0%
      Failures due to timeout: 0
      Failures due to disconnect: 0
      Failures due to no connection: 0
      Failures due to sequence error: 0
      Failures due to internal error: 0
      Failures due to other errors:
      Packet(s) arrived late: 0

# Display the history of HTTP tests.
[DeviceA] display nqa history admin test
  NQA entry (admin admin, tag test) history record(s):
    Index      Response      Status      Time
    1          64           Succeeded   2011-01-22 10:12:47.9

```

UDP jitter test configuration example

Network requirements

As shown in [Figure 45](#), configure NQA UDP jitter tests to test the delay jitter of packet transmission between Device A and Device B.

Figure 45 Network diagram



Configuration procedure

Before you make the configuration, make sure the devices can reach each other.

1. Configure Device B:

Enable the NQA server, and configure a listening service to listen to IP address 10.2.2.2 and UDP port 9000.

```
<DeviceB> system-view
[DeviceB] nqa server enable
[DeviceB] nqa server udp-echo 10.2.2.2 9000
```

2. Configure Device A:

Create a UDP jitter test group.

```
<DeviceA> system-view
[DeviceA] nqa entry admin test
[DeviceA-nqa-admin-test] type udp-jitter
```

Configure UDP jitter packets to use 10.2.2.2 as the destination IP address and port 9000 as the destination port.

```
[DeviceA-nqa-admin-test-udp-jitter] destination ip 10.2.2.2
[DeviceA-nqa-admin-test-udp-jitter] destination port 9000
```

Configure the device to perform UDP jitter tests at an interval of 1000 milliseconds.

```
[DeviceA-nqa-admin-test-udp-jitter] frequency 1000
[DeviceA-nqa-admin-test-udp-jitter] quit
```

Start UDP jitter tests.

```
[DeviceA] nqa schedule admin test start-time now lifetime forever
```

Stop UDP jitter tests after a period of time.

```
[DeviceA] undo nqa schedule admin test
```

Display the result of the last UDP jitter test.

```
[DeviceA] display nqa result admin test
```

```
NQA entry (admin admin, tag test) test results:
```

```
Destination IP address: 10.2.2.2
```

```
Send operation times: 10
```

```
Receive response times: 10
```

```
Min/Max/Average round trip time: 15/32/17
```

```
Square-Sum of round trip time: 3235
```

```
Last succeeded probe time: 2011-01-29 13:56:17.6
```

```
Extended results:
```

```
Packet loss in test: 0%
```

```
Failures due to timeout: 0
```

```
Failures due to disconnect: 0
```

```
Failures due to no connection: 0
```

```
Failures due to sequence error: 0
```

```
Failures due to internal error: 0
```

```
Failures due to other errors: 0
```

```
Packet(s) arrived late: 0
```

```
UDP-jitter results:
```

```
RTT number: 10
```

```
Min positive SD: 4
```

```
Min positive DS: 1
```

```
Max positive SD: 21
```

```
Max positive DS: 28
```

```
Positive SD number: 5
```

```
Positive DS number: 4
```

```

Positive SD sum: 52
Positive SD average: 10
Positive SD square sum: 754
Min negative SD: 1
Max negative SD: 13
Negative SD number: 4
Negative SD sum: 38
Negative SD average: 10
Negative SD square sum: 460
Positive DS sum: 38
Positive DS average: 10
Positive DS square sum: 460
Min negative DS: 6
Max negative DS: 22
Negative DS number: 5
Negative DS sum: 52
Negative DS average: 10
Negative DS square sum: 754
One way results:
Max SD delay: 15
Min SD delay: 7
Number of SD delay: 10
Sum of SD delay: 78
Square sum of SD delay: 666
SD lost packet(s): 0
Lost packet(s) for unknown reason: 0
Max DS delay: 16
Min DS delay: 7
Number of DS delay: 10
Sum of DS delay: 85
Square sum of DS delay: 787
DS lost packet(s): 0

```

Display the statistics of UDP jitter tests.

[DeviceA] display nqa statistics admin test

NQA entry (admin admin, tag test) test statistics:

```

NO. : 1
Destination IP address: 10.2.2.2
Start time: 2011-01-29 13:56:14.0
Life time: 47 seconds
Send operation times: 410          Receive response times: 410
Min/Max/Average round trip time: 1/93/19
Square-Sum of round trip time: 206176

```

Extended results:

```

Packet loss in test: 0%
Failures due to timeout: 0
Failures due to disconnect: 0
Failures due to no connection: 0
Failures due to sequence error: 0
Failures due to internal error: 0
Failures due to other errors: 0
Packet(s) arrived late: 0

```

UDP-jitter results:

```

RTT number: 410
Min positive SD: 3
Max positive SD: 30
Positive SD number: 186
Positive SD sum: 2602
Positive SD average: 13
Positive SD square sum: 45304
Min negative SD: 1
Max negative SD: 30
Negative SD number: 181
Negative SD sum: 181
Min positive DS: 1
Max positive DS: 79
Positive DS number: 158
Positive DS sum: 1928
Positive DS average: 12
Positive DS square sum: 31682
Min negative DS: 1
Max negative DS: 78
Negative DS number: 209
Negative DS sum: 209

```

```

Negative SD average: 13
Negative SD square sum: 46994
One way results:
Max SD delay: 46
Min SD delay: 7
Number of SD delay: 410
Sum of SD delay: 3705
Square sum of SD delay: 45987
SD lost packet(s): 0
Lost packet(s) for unknown reason: 0

Negative DS average: 14
Negative DS square sum: 3030
Max DS delay: 46
Min DS delay: 7
Number of DS delay: 410
Sum of DS delay: 3891
Square sum of DS delay: 49393
DS lost packet(s): 0

```

NOTE:

The **display nqa history** command does not show the results of UDP jitter tests. To know the result of a UDP jitter test, use the **display nqa result** command to view the probe results of the latest NQA test, or use the **display nqa statistics** command to view the statistics of NQA tests.

SNMP test configuration example

Network requirements

As shown in [Figure 46](#), configure NQA SNMP tests to test the time it takes for Device A to send an SNMP query packet to the SNMP agent and receive a response packet.

Figure 46 Network diagram



Configuration procedure

Before you make the configuration, make sure the devices can reach each other.

1. Configure the SNMP agent (Device B):

Enable the SNMP agent service and set the SNMP version to **all**, the read community to **public**, and the write community to **private**.

```

<DeviceB> system-view
[DeviceB] snmp-agent sys-info version all
[DeviceB] snmp-agent community read public
[DeviceB] snmp-agent community write private

```

2. Configure Device A:

Create an SNMP test group, and configure SNMP packets to use 10.2.2.2 as their destination IP address.

```

<DeviceA> system-view
[DeviceA] nqa entry admin test
[DeviceA-nqa-admin-test] type snmp
[DeviceA-nqa-admin-test-snmp] destination ip 10.2.2.2

```

Enable the saving of history records.

```

[DeviceA-nqa-admin-test-snmp] history-record enable

```

```

[DeviceA-nqa-admin-test-snmp] quit
# Start SNMP tests.
[DeviceA] nqa schedule admin test start-time now lifetime forever
# Stop the SNMP tests after a period of time.
[DeviceA] undo nqa schedule admin test
# Display the results of the last SNMP test.
[DeviceA] display nqa result admin test
  NQA entry (admin admin, tag test) test results:
    Destination IP address: 10.2.2.2
      Send operation times: 1                Receive response times: 1
      Min/Max/Average round trip time: 50/50/50
      Square-Sum of round trip time: 2500
      Last succeeded probe time: 2011-01-22 10:24:41.1
    Extended results:
      Packet loss in test: 0%
      Failures due to timeout: 0
      Failures due to disconnect: 0
      Failures due to no connection: 0
      Failures due to sequence error: 0
      Failures due to internal error: 0
      Failures due to other errors: 0
      Packet(s) arrived late: 0
# Display the history of SNMP tests.
[DeviceA] display nqa history admin test
  NQA entry (admin admin, tag test) history record(s):
    Index      Response      Status      Time
    ---      -
    1          50           Timeout    2011-01-22 10:24:41.1

```

TCP test configuration example

Network requirements

As shown in [Figure 47](#), configure NQA TCP tests to test the time for establishing a TCP connection between Device A and Device B.

Figure 47 Network diagram



Configuration procedure

Before you make the configuration, make sure the devices can reach each other.

1. Configure Device B:

```
# Enable the NQA server, and configure a listening service to listen to IP address 10.2.2.2 and TCP port 9000.
```

```
<DeviceB> system-view
[DeviceB] nqa server enable
```

```
[DeviceB] nqa server tcp-connect 10.2.2.2 9000
```

2. Configure Device A:

Create a TCP test group.

```
<DeviceA> system-view
```

```
[DeviceA] nqa entry admin test
```

```
[DeviceA-nqa-admin-test] type tcp
```

Configure TCP probe packets to use 10.2.2.2 as the destination IP address and port 9000 as the destination port.

```
[DeviceA-nqa-admin-test-tcp] destination ip 10.2.2.2
```

```
[DeviceA-nqa-admin-test-tcp] destination port 9000
```

Enable the saving of history records.

```
[DeviceA-nqa-admin-test-tcp] history-record enable
```

```
[DeviceA-nqa-admin-test-tcp] quit
```

Start TCP tests.

```
[DeviceA] nqa schedule admin test start-time now lifetime forever
```

Stop the TCP tests after a period of time.

```
[DeviceA] undo nqa schedule admin test
```

Display the results of the last TCP test.

```
[DeviceA] display nqa result admin test
```

```
NQA entry (admin admin, tag test) test results:
```

```
Destination IP address: 10.2.2.2
```

```
Send operation times: 1
```

```
Receive response times: 1
```

```
Min/Max/Average round trip time: 13/13/13
```

```
Square-Sum of round trip time: 169
```

```
Last succeeded probe time: 2011-01-22 10:27:25.1
```

```
Extended results:
```

```
Packet loss in test: 0%
```

```
Failures due to timeout: 0
```

```
Failures due to disconnect: 0
```

```
Failures due to no connection: 0
```

```
Failures due to sequence error: 0
```

```
Failures due to internal error: 0
```

```
Failures due to other errors: 0
```

```
Packet(s) arrived late: 0
```

Display the history of TCP tests.

```
[DeviceA] display nqa history admin test
```

```
NQA entry (admin admin, tag test) history record(s):
```

| Index | Response | Status | Time |
|-------|----------|-----------|-----------------------|
| 1 | 13 | Succeeded | 2011-01-22 10:27:25.1 |

UDP echo test configuration example

Network requirements

As shown in [Figure 48](#), configure NQA UDP echo tests to test the round-trip time between Device A and Device B. The destination port number is 8000.

Figure 48 Network diagram



Configuration procedure

Before you make the configuration, make sure the devices can reach each other.

1. Configure Device B:

Enable the NQA server, and configure a listening service to listen to IP address 10.2.2.2 and UDP port 8000.

```
<DeviceB> system-view
[DeviceB] nqa server enable
[DeviceB] nqa server udp-echo 10.2.2.2 8000
```

2. Configure Device A:

Create a UDP echo test group.

```
<DeviceA> system-view
[DeviceA] nqa entry admin test
[DeviceA-nqa-admin-test] type udp-echo
# Configure UDP packets to use 10.2.2.2 as the destination IP address and port 8000 as the destination port.
[DeviceA-nqa-admin-test-udp-echo] destination ip 10.2.2.2
[DeviceA-nqa-admin-test-udp-echo] destination port 8000
```

Enable the saving of history records.

```
[DeviceA-nqa-admin-test-udp-echo] history-record enable
[DeviceA-nqa-admin-test-udp-echo] quit
```

Start UDP echo tests.

```
[DeviceA] nqa schedule admin test start-time now lifetime forever
```

Stop UDP echo tests after a period of time.

```
[DeviceA] undo nqa schedule admin test
```

Display the results of the last UDP echo test.

```
[DeviceA] display nqa result admin test
NQA entry (admin admin, tag test) test results:
  Destination IP address: 10.2.2.2
    Send operation times: 1                Receive response times: 1
  Min/Max/Average round trip time: 25/25/25
  Square-Sum of round trip time: 625
  Last succeeded probe time: 2011-01-22 10:36:17.9
Extended results:
  Packet loss in test: 0%
  Failures due to timeout: 0
  Failures due to disconnect: 0
  Failures due to no connection: 0
  Failures due to sequence error: 0
  Failures due to internal error: 0
```

```

Failures due to other errors: 0
Packet(s) arrived late: 0
# Display the history of UDP echo tests.
[DeviceA] display nqa history admin test
NQA entry (admin admin, tag test) history record(s):
  Index      Response      Status      Time
  1          25            Succeeded   2011-01-22 10:36:17.9

```

Voice test configuration example

Network requirements

As shown in [Figure 49](#), configure NQA voice tests to test the delay jitter of voice packet transmission and voice quality between Device A and Device B.

Figure 49 Network diagram



Configuration procedure

Before you make the configuration, make sure the devices can reach each other.

1. Configure Device B:

Enable the NQA server, and configure a listening service to listen to IP address 10.2.2.2 and UDP port 9000.

```

<DeviceB> system-view
[DeviceB] nqa server enable
[DeviceB] nqa server udp-echo 10.2.2.2 9000

```

2. Configure Device A:

Create a voice test group.

```

<DeviceA> system-view
[DeviceA] nqa entry admin test
[DeviceA-nqa-admin-test] type voice

```

Configure voice probe packets to use 10.2.2.2 as the destination IP address and port 9000 as the destination port.

```

[DeviceA-nqa-admin-test-voice] destination ip 10.2.2.2
[DeviceA-nqa-admin-test-voice] destination port 9000
[DeviceA-nqa-admin-test-voice] quit

```

Start voice tests.

```

[DeviceA] nqa schedule admin test start-time now lifetime forever

```

Stop the voice tests after a period of time.

```

[DeviceA] undo nqa schedule admin test

```

Display the result of the last voice test.

```

[DeviceA] display nqa result admin test
NQA entry (admin admin, tag test) test results:
  Destination IP address: 10.2.2.2

```


Send operation times: 1000 Receive response times: 1000
Min/Max/Average round trip time: 31/1328/33
Square-Sum of round trip time: 2844813
Last succeeded probe time: 2011-01-13 09:49:31.1

Extended results:

Packet loss in test: 0%
Failures due to timeout: 0
Failures due to disconnect: 0
Failures due to no connection: 0
Failures due to sequence error: 0
Failures due to internal error: 0
Failures due to other errors: 0
Packet(s) arrived late: 0

Voice results:

RTT number: 1000

| | |
|-------------------------------|---------------------------------|
| Min positive SD: 1 | Min positive DS: 1 |
| Max positive SD: 204 | Max positive DS: 1297 |
| Positive SD number: 257 | Positive DS number: 259 |
| Positive SD sum: 759 | Positive DS sum: 1797 |
| Positive SD average: 2 | Positive DS average: 6 |
| Positive SD square sum: 54127 | Positive DS square sum: 1691967 |
| Min negative SD: 1 | Min negative DS: 1 |
| Max negative SD: 203 | Max negative DS: 1297 |
| Negative SD number: 255 | Negative DS number: 259 |
| Negative SD sum: 759 | Negative DS sum: 1796 |
| Negative SD average: 2 | Negative DS average: 6 |
| Negative SD square sum: 53655 | Negative DS square sum: 1691776 |

One way results:

| | |
|--------------------------------------|--------------------------------|
| Max SD delay: 343 | Max DS delay: 985 |
| Min SD delay: 343 | Min DS delay: 985 |
| Number of SD delay: 1 | Number of DS delay: 1 |
| Sum of SD delay: 343 | Sum of DS delay: 985 |
| Square sum of SD delay: 117649 | Square sum of DS delay: 970225 |
| SD lost packet(s): 0 | DS lost packet(s): 0 |
| Lost packet(s) for unknown reason: 0 | |

Voice scores:

MOS value: 4.38 ICPIF value: 0

Display the statistics of voice tests.

[DeviceA] display nqa statistics admin test

NQA entry (admin admin, tag test) test statistics:

NO. : 1

Destination IP address: 10.2.2.2

Start time: 2011-01-13 09:45:37.8

Life time: 331 seconds

Send operation times: 4000 Receive response times: 4000

Min/Max/Average round trip time: 15/1328/32

Square-Sum of round trip time: 7160528

Extended results:

```

Packet loss in test: 0%
Failures due to timeout: 0
Failures due to disconnect: 0
Failures due to no connection: 0
Failures due to sequence error: 0
Failures due to internal error: 0
Failures due to other errors: 0
Packet(s) arrived late: 0
Voice results:
RTT number: 4000
Min positive SD: 1
Max positive SD: 360
Positive SD number: 1030
Positive SD sum: 4363
Positive SD average: 4
Positive SD square sum: 497725
Min negative SD: 1
Max negative SD: 360
Negative SD number: 1028
Negative SD sum: 1028
Negative SD average: 4
Negative SD square sum: 495901
Min positive DS: 1
Max positive DS: 1297
Positive DS number: 1024
Positive DS sum: 5423
Positive DS average: 5
Positive DS square sum: 2254957
Min negative DS: 1
Max negative DS: 1297
Negative DS number: 1022
Negative DS sum: 1022
Negative DS average: 5
Negative DS square sum: 5419
One way results:
Max SD delay: 359
Min SD delay: 0
Number of SD delay: 4
Sum of SD delay: 1390
Square sum of SD delay: 483202
SD lost packet(s): 0
Lost packet(s) for unknown reason: 0
Max DS delay: 985
Min DS delay: 0
Number of DS delay: 4
Sum of DS delay: 1079
Square sum of DS delay: 973651
DS lost packet(s): 0
Voice scores:
Max MOS value: 4.38
Min MOS value: 4.38
Max ICPIF value: 0
Min ICPIF value: 0

```

NOTE:

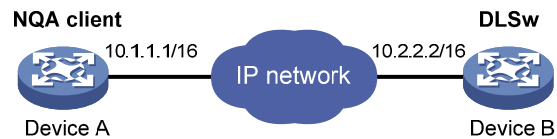
The **display nqa history** command cannot show you the results of voice tests. To know the result of a voice test, use the **display nqa result** command to view the probe results of the latest NQA test, or use the **display nqa statistics** command to view the statistics of NQA tests.

DLSw test configuration example

Network requirements

As shown in [Figure 50](#), configure NQA DLSw tests to test the response time of the DLSw device.

Figure 50 Network diagram



Configuration procedure

Before you make the configuration, make sure the devices can reach each other.

Create a DLSw test group, and configure DLSw probe packets to use 10.2.2.2 as the destination IP address.

```
<DeviceA> system-view
[DeviceA] nqa entry admin test
[DeviceA-nqa-admin-test] type dlsw
[DeviceA-nqa-admin-test-dlsw] destination ip 10.2.2.2
```

Enable the saving of history records.

```
[DeviceA-nqa-admin-test-dlsw] history-record enable
[DeviceA-nqa-admin-test-dlsw] quit
```

Start DLSw tests.

```
[DeviceA] nqa schedule admin test start-time now lifetime forever
```

Stop the DLSw tests after a period of time.

```
[DeviceA] undo nqa schedule admin test
```

Display the result of the last DLSw test.

```
[DeviceA] display nqa result admin test
NQA entry (admin admin, tag test) test results:
  Destination IP address: 10.2.2.2
  Send operation times: 1          Receive response times: 1
  Min/Max/Average round trip time: 19/19/19
  Square-Sum of round trip time: 361
  Last succeeded probe time: 2011-01-22 10:40:27.7
Extended results:
  Packet loss in test: 0%
  Failures due to timeout: 0
  Failures due to disconnect: 0
  Failures due to no connection: 0
  Failures due to sequence error: 0
  Failures due to internal error: 0
  Failures due to other errors: 0
  Packet(s) arrived late: 0
```

Display the history of DLSw tests.

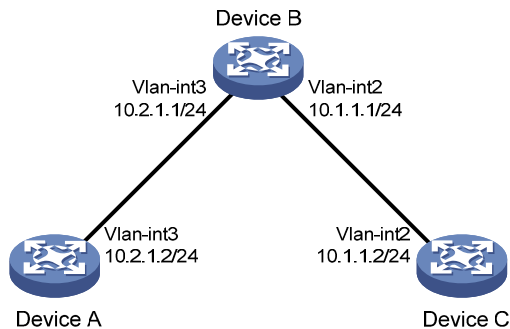
```
[DeviceA] display nqa history admin test
NQA entry (admin admin, tag test) history record(s):
  Index      Response      Status      Time
  ---      -
  1          19            Succeeded   2011-01-22 10:40:27.7
```

NQA collaboration configuration example

Network requirements

As shown in [Figure 51](#), configure a static route to Device C on Device A, with Device B as the next hop. Associate the static route, track entry, and NQA test group to verify whether static route is active in real time.

Figure 51 Network diagram



Configuration procedure

1. Assign each interface an IP address. (Details not shown.)
2. On Device A, configure a unicast static route and associate the static route with a track entry.
Configure a static route, whose destination address is 10.2.1.1, and associate the static route with track entry 1.

```
<DeviceA> system-view
[DeviceA] ip route-static 10.1.1.2 24 10.2.1.1 track 1
```
3. On Device A, create an NQA test group:
Create an NQA test group with the administrator name being **admin** and operation tag being **test**.

```
[DeviceA] nqa entry admin test
```


Configure the test type of the NQA test group as ICMP echo.

```
[DeviceA-nqa-admin-test] type icmp-echo
```


Configure ICMP echo requests to use 10.2.1.1 as their destination IP address.

```
[DeviceA-nqa-admin-test-icmp-echo] destination ip 10.2.1.1
```


Configure the device to perform tests at an interval of 100 milliseconds.

```
[DeviceA-nqa-admin-test-icmp-echo] frequency 100
```


Create reaction entry 1. If the number of consecutive probe failures reaches 5, collaboration with other modules is triggered.

```
[DeviceA-nqa-admin-test-icmp-echo] reaction 1 checked-element probe-fail
threshold-type consecutive 5 action-type trigger-only
[DeviceA-nqa-admin-test-icmp-echo] quit
```


Configure the test start time and test duration for the test group.

```
[DeviceA] nqa schedule admin test start-time now lifetime forever
```
4. On Device A, create the track entry:
Create track entry 1, and associate it with reaction entry 1 of the NQA test group (admin-test).

```
[DeviceA] track 1 nqa entry admin test reaction 1
```

Verifying the configuration

On Device A, display information about all the track entries.

```
[DeviceA] display track all
Track ID: 1
  Status: Positive
  Notification delay: Positive 0, Negative 0 (in seconds)
  Reference object:
    NQA entry: admin test
    Reaction: 1
```

Display brief information about active routes in the routing table on Device A.

```
[DeviceA] display ip routing-table
Routing Tables: Public
          Destinations : 5          Routes : 5
```

| Destination/Mask | Proto | Pre | Cost | NextHop | Interface |
|------------------|--------|-----|------|-----------|-----------|
| 10.1.1.0/24 | Static | 60 | 0 | 10.2.1.1 | Vlan3 |
| 10.2.1.0/24 | Direct | 0 | 0 | 10.2.1.2 | Vlan3 |
| 10.2.1.2/32 | Direct | 0 | 0 | 127.0.0.1 | InLoop0 |
| 127.0.0.0/8 | Direct | 0 | 0 | 127.0.0.1 | InLoop0 |
| 127.0.0.1/32 | Direct | 0 | 0 | 127.0.0.1 | InLoop0 |

The output shows that the static route with the next hop 10.2.1.1 is active, and the status of the track entry is positive. The static route configuration works.

Remove the IP address of VLAN-interface 3 on Device B.

```
<DeviceB> system-view
[DeviceB] interface vlan-interface 3
[DeviceB-Vlan-interface3] undo ip address
```

On Device A, display information about all the track entries.

```
[DeviceA] display track all
Track ID: 1
  Status: Negative
  Notification delay: Positive 0, Negative 0 (in seconds)
  Reference object:
    NQA entry: admin test
    Reaction: 1
```

Display brief information about active routes in the routing table on Device A.

```
[DeviceA] display ip routing-table
Routing Tables: Public
          Destinations : 4          Routes : 4
```

| Destination/Mask | Proto | Pre | Cost | NextHop | Interface |
|------------------|--------|-----|------|-----------|-----------|
| 10.2.1.0/24 | Direct | 0 | 0 | 10.2.1.2 | Vlan3 |
| 10.2.1.2/32 | Direct | 0 | 0 | 127.0.0.1 | InLoop0 |
| 127.0.0.0/8 | Direct | 0 | 0 | 127.0.0.1 | InLoop0 |
| 127.0.0.1/32 | Direct | 0 | 0 | 127.0.0.1 | InLoop0 |

The output shows that the next hop 10.2.1.1 of the static route is not reachable, and the status of the track entry is negative. The static route does not work.

Configuring sFlow

Both bridge mode (Layer 2) and route mode (Layer 3) Ethernet ports support sFlow. The term "interface" in this chapter collectively refers to these two types of ports. You can use the **port link-mode** command to set an Ethernet port to operate in bridge or route mode (see *Layer 2—LAN Switching Configuration Guide*).

sFlow overview

Introduction to sFlow

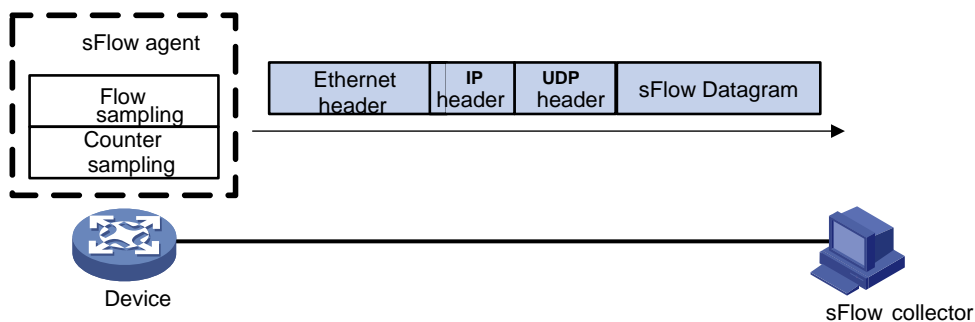
Sampled Flow (sFlow) is a traffic monitoring technology mainly used to collect and analyze traffic statistics.

As shown in [Figure 52](#), the sFlow system involves an sFlow agent embedded in a device and a remote sFlow collector. The sFlow agent collects traffic statistics and packet information from the sFlow-enabled interfaces and encapsulates them into sFlow packets. When the sFlow packet buffer is full, or the age time of sFlow packets is reached, (the age time is one second), the sFlow agent sends the packets to a specified sFlow collector. The sFlow collector analyzes the sFlow packets and displays the results.

sFlow has the following two sampling mechanisms:

- Flow sampling: Packet-based sampling, used to obtain packet content information.
- Counter sampling: Time-based sampling, used to obtain port traffic statistics.

Figure 52 sFlow system



As a traffic monitoring technology, sFlow has the following advantages:

- Supporting traffic monitoring on Gigabit and higher-speed networks.
- Providing good scalability to allow one sFlow collector to monitor multiple sFlow agents.
- Saving cost by embedding the sFlow agent in a device, instead of using a dedicated sFlow agent device.

NOTE:

Only the sFlow agent function is supported on the switch.

sFlow operation

sFlow operates in the following ways:

1. Before enabling the sFlow function, configure the sFlow agent and sFlow collector on the device.
2. With flow sampling enabled on an Ethernet interface, the sFlow agent samples packets and encapsulates them into sFlow packets. For the configuration, see "[Configuring flow sampling](#)."
3. With counter sampling enabled on an Ethernet interface, the sFlow agent periodically collects the statistics of the interface and encapsulates the statistics into sFlow packets. For the configuration, see "[Configuring flow sampling](#)."

Configuring sFlow

Before configuring sFlow, complete the following tasks:

- Configure the IP address, flow sampling, and counter sampling of the sFlow collector on the device.
- Configure the sFlow collector.

Configuring the sFlow agent and sFlow collector

The sFlow feature enables the remote sFlow collector to monitor the network and analyze sFlow packet statistics.

To configure the sFlow agent and sFlow collector:

| Step | Command | Remarks |
|---|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Specify the IP address for the sFlow agent. | sflow agent { ip <i>ip-address</i> ipv6 <i>ipv6-address</i> } | Optional
Not specified by default. The device periodically checks the existence of the sFlow agent address. If the sFlow agent has no IP address configured, the device automatically selects an interface IP address for the sFlow agent but does not save the selected IP address.
NOTE: <ul style="list-style-type: none">• HP recommends configuring an IP address manually for the sFlow agent.• Only one IP address can be specified for the sFlow agent on the device. |
| 3. Configure the sFlow collector. | sflow collector <i>collector-id</i> { { ip <i>ip-address</i> ipv6 <i>ipv6-address</i> } datagram-size <i>size</i> description <i>text</i> port <i>port-number</i> time-out <i>seconds</i> } * | By default, the device presets a number of sFlow collectors.
Use the display sflow command to display the parameters of the preset sFlow collectors. |
| 4. Specify the source IP address of sent sFlow packets. | sflow source { ip <i>ip-address</i> ipv6 <i>ipv6-address</i> } * | Optional
Not specified by default. |

Configuring flow sampling

| Step | Command | Remarks |
|---|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter Layer 2 Ethernet interface view/Layer 3 Ethernet interface view. | interface <i>interface-type</i>
<i>interface-number</i> | Only the HP 5500 EI switch series supports this command on Layer 3 Ethernet interfaces. |
| 3. Set the Flow sampling mode. | sflow sampling-mode
{ determine random } | Optional
random by default. |
| 4. Set the interval for flow sampling. | sflow sampling-rate <i>interval</i> | Not set by default. |
| 5. Set the maximum copied length of a sampled packet. | sflow flow max-header <i>length</i> | Optional
By default, up to 128 bytes of a sampled packet can be copied. You are recommended to use the default value. |
| 6. Specify the sFlow collector for flow sampling. | sflow flow collector
<i>collector-id</i> | No collector is specified for flow sampling by default. |

NOTE:

The switch does not support the flow sampling mode **determine**.

Configuring counter sampling

| Step | Command | Remarks |
|--|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter Layer 2 interface view/Layer 3 Ethernet interface view. | interface <i>interface-type</i>
<i>interface-number</i> | Only the HP 5500 EI switch series supports Layer 3 Ethernet interface view. |
| 3. Set the interval for counter sampling. | sflow counter interval <i>seconds</i> | Counter sampling is disabled by default. |
| 4. Specify the sFlow collector for counter sampling. | sflow counter collector
<i>collector-id</i> | No collector is specified for counter sampling by default. |

Displaying and maintaining sFlow

| Task | Command | Remarks |
|--|--|-----------------------|
| Display sFlow configuration information. | display sflow [slot <i>slot-number</i>] [
{ begin exclude include }
<i>regular-expression</i>] | Available in any view |

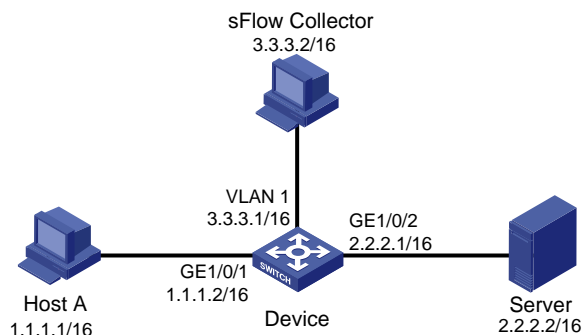
sFlow configuration example

Network requirements

As shown in [Figure 53](#), Host A is connected with the server through the device (sFlow agent).

Enable sFlow (including flow sampling and counter sampling) on GigabitEthernet 1/0/1 to monitor traffic on the port. The device sends sFlow packets through GigabitEthernet 1/0/3 to the sFlow collector, which analyzes the sFlow packets and displays results.

Figure 53 Network diagram



Configuration procedure

1. Configure the sFlow agent and sFlow collector:

Configure the IP address of vlan-interface 1 on Device as 3.3.3.1/16.

```
<Device> system-view
[Device] interface vlan-interface 1
[Device-Vlan-interface1] ip address 3.3.3.1 16
[Device-Vlan-interface1] quit
```

Specify the IP address for the sFlow agent.

```
[Device] sflow agent ip 3.3.3.1
```

Specify sFlow collector ID 2, IP address 3.3.3.2, the default port number, and description of **netserver** for the sFlow collector.

```
[Device] sflow collector 2 ip 3.3.3.2 description netserver
```

2. Configure counter sampling:

Set the counter sampling interval to 120 seconds.

```
[Device] interface gigabitethernet 1/0/1
[Device-GigabitEthernet1/0/1] sflow counter interval 120
```

Specify sFlow collector 2 for counter sampling.

```
[Device-GigabitEthernet1/0/1] sflow counter collector 2
```

3. Configure flow sampling:

Set the Flow sampling mode and sampling interval.

```
[Device-GigabitEthernet1/0/1] sflow sampling-mode random
[Device-GigabitEthernet1/0/1] sflow sampling-rate 4000
```

Specify sFlow collector 2 for flow sampling.

```
[Device-GigabitEthernet1/0/1] sflow flow collector 2
```

```
# Display the sFlow configuration and operation information.
[Device-GigabitEthernet1/0/1] display sflow
sFlow Version: 5
sFlow Global Information:
Agent          IP:3.3.3.1
Collector Information:
ID   IP                               Port  Aging  Size  Description
1    3.3.3.2                             6343  0      1400  netserver
2    3.3.3.2                             6543  N/A    1400
3    3.3.3.2                             6343  0      1400
4    3.3.3.2                             6343  0      1400
5    3.3.3.2                             6343  0      1400
6    3.3.3.2                             6343  0      1400
7    3.3.3.2                             6343  0      1400
8    3.3.3.2                             6343  0      1400
9    3.3.3.2                             6343  0      1400
10   3.3.3.2                             6343  0      1400
sFlow Port Information:
Interface CID  Interval(s) FID  MaxHLen  Rate  Mode  Status
GigabitEthernet1/0/1  2    120    2    128    4000  Random  Active
```

The output shows that GigabitEthernet 1/0/1 enabled with sFlow is active, the counter sampling interval is 120 seconds, the Flow sampling interval is 4000, all of which indicate sFlow operates normally.

Troubleshooting sFlow configuration

Symptom

The remote sFlow collector cannot receive sFlow packets.

Analysis

- The sFlow collector has no IP address specified.
- No interface is enabled with sFlow to sample data.
- The IP address of the sFlow collector specified on the sFlow agent is different from that of the remote sFlow collector.
- No IP address is configured for the Layer 3 interface on the device, or the IP address is configured, but the UDP packets with the IP address being the source cannot reach the sFlow collector.
- The physical link between the device and the sFlow collector fails.

Solution

1. Check whether sFlow is correctly configured by displaying sFlow configuration with the **display sflow** command.
2. Check whether the correct IP address is configured for the device to communicate with the sFlow collector.
3. Check whether the physical link between the device and the sFlow collector is normal.

Configuring IPC

This chapter provides an overview of IPC and describes the IPC monitoring commands.

Overview

Inter-Process Communication (IPC) provides a reliable communication mechanism among processing units, typically CPUs. IPC is typically used on a distributed device or in an IRF fabric to provide reliable inter-card or inter-device transmission. This section describes the basic IPC concepts.

Node

An IPC node is an independent IPC-capable processing unit, typically, a CPU.

This series of devices are centralized devices that have only one CPU. The IRF fabrics formed by them have multiple CPUs, or IPC nodes.

Link

An IPC link is a connection between any two IPC nodes. Any two IPC nodes have one and only one IPC link for sending and receiving packets. All IPC nodes are fully meshed.

IPC links are created when the system is initialized. An IPC node, upon startup, sends handshake packets to other nodes. If the handshake succeeds, a connection is established.

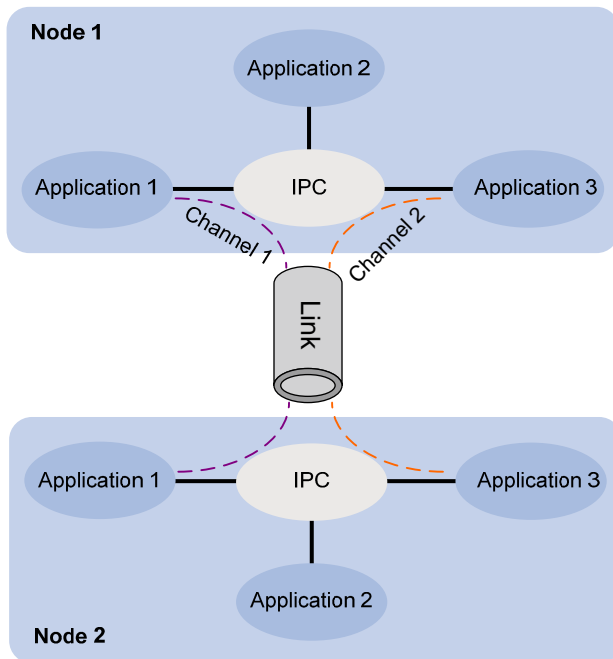
The system uses link status to identify the link connectivity between two nodes. An IPC node can have multiple links, and each link has its own status.

Channel

A channel is the communication interface between peer upper layer application modules that use different IPC nodes. Each node assigns a locally unique channel number to each upper layer application module for identification.

An upper layer application module sends data to an IPC module in a channel, and the IPC module sends the data to a peer node across a link, as shown in [Figure 54](#).

Figure 54 Relationship between a node, link and channel



Packet sending modes

IPC uses one of the following modes to send packets for upper layer application modules:

- **Unicast**—One node sends packets to another node.
- **Multicast**—One node sends packets to multiple nodes. This mode includes broadcast, a special multicast. To use multicast mode, an application module must create a multicast group that includes a set of nodes. Multicasts destined for this group are sent to all the nodes in the group. An application module can create multiple multicast groups. Creation and deletion of a multicast group or group member depend on the application module.
- **Mixcast**—Supports both unicast and multicast.

IPC assigns one queue for each mode. An upper layer application module automatically selects one mode as needed.

Enabling IPC performance statistics

The IPC performance statistics function provides the most recent 10-second, 1-minute, and 5-minute traffic input and output statistics for IPC nodes. If this function is disabled, the **display ipc performance** command displays the statistics collected before IPC performance statistics was disabled.

To enable IPC performance statistics:

| Task | Command | Remarks |
|---|--|---------------------------------------|
| Enable IPC performance statistics in user view. | ipc performance enable { node <i>node-id</i> self-node } [channel <i>channel-id</i>] | By default, the function is disabled. |

Displaying and maintaining IPC

| Task | Command | Remarks |
|---|--|------------------------|
| Display IPC node information. | display ipc node [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display channel information for a node. | display ipc channel { node <i>node-id</i> self-node } [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display queue information for a node. | display ipc queue { node <i>node-id</i> self-node } [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display multicast group information for a node. | display ipc multicast-group { node <i>node-id</i> self-node } [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display packet information for a node. | display ipc packet { node <i>node-id</i> self-node } [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display link status information for a node. | display ipc link { node <i>node-id</i> self-node } [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display IPC performance statistics of a node. | display ipc performance { node <i>node-id</i> self-node } [channel <i>channel-id</i>] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Clear IPC performance statistics of a node. | reset ipc performance [node <i>node-id</i> self-node] [channel <i>channel-id</i>] | Available in user view |

Configuring PoE

Overview

Power over Ethernet (PoE) enables a power sourcing equipment (PSE) to supply power to powered devices (PDs) through Ethernet interfaces over twisted pair cables.

Advantages

- **Reliable**—Power is supplied in a centralized way so that it is very convenient to provide a backup power supply.
- **Easy to connect**—A network terminal requires no external power supply but only an Ethernet cable.
- **Standard**—In compliance with IEEE 802.3af, and a globally uniform power interface is adopted.
- **Promising**—It can be applied to IP telephones, wireless LAN access points (APs), portable chargers, card readers, web cameras, and data collectors.

PoE concepts

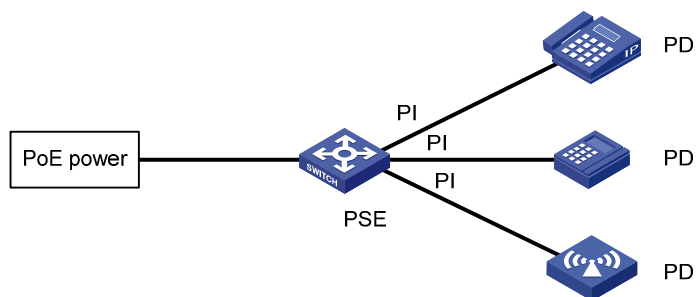
As shown in [Figure 55](#), a PoE system comprises PoE power, PSE, power interface (PI), and PD:

- **PoE power**—The whole PoE system is powered by the PoE power.
- **PSE**—A PSE supplies power for PDs. A PSE can examine the Ethernet cables connected to PoE interfaces, search for PDs, classify them, and supply power to them. When detecting that a PD is unplugged, the PSE stops supplying power to the PD. A PSE can be built-in (Endpoint) or external (Midspan). The switch uses built-in PSEs. To display the mapping between a PSE ID and the slot number of an interface card, execute the **display poe device** command.

The PSE ID is the *switch member ID* $\times 3 + 1$. For example, if the member ID of the device is 3, the PSE ID of the device is $3 \times 3 + 1 = 10$.

- **PI**—An Ethernet interface with the PoE capability is called PoE interface.
- **PD**—A PD accepts power from the PSE, including IP phones, wireless APs, chargers of portable devices, POS, and web cameras. The PD that is being powered by the PSE can be connected to another power supply unit for redundancy power backup.

Figure 55 PoE system diagram



Protocol specification

The protocol specification related to PoE is IEEE 802.3af.

PoE configuration task list

You can configure a PoE interface by using either of the following methods:

- At the command line interface (CLI).
- Through configuring the PoE profile and applying the PoE profile to the PoE interface.

To configure a single PoE interface, configure it at the CLI. To configure PoE interfaces in batches, use the PoE profile. For a PoE configuration parameter of a PoE interface, you can only select one mode (including modification and removal of a PoE interface).

Configuration guidelines

- Before configuring PoE, make sure the PoE power supply and PSE are operating normally. Otherwise, you cannot configure PoE or the configured PoE function does not take effect.
- Turning off the PoE power supply during the startup of the device might cause the PoE configuration in the PoE profile invalid.

Complete these tasks to configure PoE:

| Task | Remarks |
|---|---|
| Enabling PoE: | |
| Enabling PoE for a PoE interface | Required. |
| Detecting PDs: | |
| Enabling the PSE to detect nonstandard PDs | Optional. |
| Configuring a PD disconnection detection mode | Optional. |
| Configuring the PoE power: | |
| Configuring the maximum PoE interface power | Optional. |
| Configuring PoE power management: | |
| Configuring PoE interface power management | Optional. |
| Configuring the PoE monitoring function: | |
| Configuring PSE power monitoring | Optional. |
| Monitoring PD | Optional.
The device automatically monitors PDs when supplying power to them, so no configuration is required. |
| Configuring PoE interface through PoE profile: | |
| Configuring PoE profile | Optional. |
| Applying PoE profile | Optional. |
| Upgrading PSE processing software in service | Optional. |

Enabling PoE

Enabling PoE for a PoE interface

The system does not supply power to or reserve power for the PDs connected to a PoE interface if the PoE interface is not enabled with the PoE function.

You are allowed to enable PoE for a PoE interface if the PoE interface will not result in PoE power overload; otherwise, whether you can enable PoE for the PoE interface depends on whether the PoE interface is enabled with the PoE power management function. For more information about the PoE interface power management function, see "[Configuring PoE interface power management](#)".

- If the PoE interface is not enabled with the PoE power management function, you are not allowed to enable PoE for the PoE interface.
- If the PoE interface is enabled with the PoE power management function, you are allowed to enable PoE for the PoE interface (whether the PDs can be powered depends on other factors, for example, the power supply priority of the PoE interface).

The PSE supplies power over category 3/5 twisted pair cable for a PoE interface in the following modes:

- **Over signal wires**—The PSE uses data pairs (pins 1, 2 and 3, 6) to supply DC power to PDs.
- **Over spare wires**—The PSE uses spare pairs (pins 4, 5 and 7, 8) to supply DC power to PDs.

Configuration guidelines

- When the sum of the power consumption of all powered PoE interfaces on a PSE exceeds the maximum power of the PSE, the system considers the PSE overloaded (The maximum PSE power is decided by the user configuration).
- A PSE can supply power to a PD only when the selected power supply mode is supported by both the PSE and PD. If the PSE and PD support different power supply modes (for example, the PSE does not support power over spare wires, while the PD supports power over spare wires), you have to change the order of the lines in the twisted pair cable to supply power to the PD.
- The switch's PoE interfaces can supply power only over signal wires.

To enable PoE for a PoE interface:

| Step | Command | Remarks |
|---|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter PoE interface view. | interface <i>interface-type</i>
<i>interface-number</i> | N/A |
| 3. Enable PoE for the PoE interface. | poe enable | By default, this function is disabled. |
| 4. Configure a description for the PD connected to the PoE interface. | poe pd-description <i>text</i> | Optional.
By default, no description for the PD connected to the PoE interface is available. |

Detecting PDs

Enabling the PSE to detect nonstandard PDs

There are standard PDs and nonstandard PDs. Usually, the PSE can detect only standard PDs and supply power to them. The PSE can detect nonstandard PDs and supply power to them only after the PSE is enabled to detect nonstandard PDs.

To enable the PSE to detect nonstandard PDs:

| Step | Command | Remarks |
|--|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enable the PSE to detect nonstandard PDs. | poe legacy enable pse <i>pse-id</i> | By default, the PSE can detect standard PDs rather than nonstandard PDs. |

Configuring a PD disconnection detection mode

To detect the PD connection with PSE, PoE provides two detection modes: AC detection and DC detection. The AC detection mode is energy saving relative to the DC detection mode.

To configure a PD disconnection detection mode:

| Step | Command | Remarks |
|---|-----------------------------------|---|
| 1. Enter system view. | system-view | N/A |
| 2. Configure a PD disconnection detection mode. | poe disconnect { ac dc } | Optional.
The default PD disconnection detection mode is AC detection. |

If you change the PD disconnection detection mode when the device is running, the connected PDs will be powered off. Therefore, be cautious to do so.

Configuring the PoE power

Configuring the maximum PoE interface power

The maximum PoE interface power is the maximum power that the PoE interface can provide to the connected PD. If the power required by the PD is larger than the maximum PoE interface power, the PoE interface will not supply power to the PD.

For HP 5500-48G-PoE+ EI Switch with 2 Interface Slots(JG240A), HP 5500-48G-PoE+ EI TAA Switch with 2 Interface Slots(JG253A), and HP 5500-48G-PoE+ SI Switch with 2 Interface Slots(JG239A), the total PoE power of ports numbered 1 through 24 is 370 W, and that of ports numbered 25 through 48 is 370 W.

To configure the maximum PSE power:

| Step | Command | Remarks |
|---|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter PoE interface view. | interface <i>interface-type</i>
<i>interface-number</i> | N/A |
| 3. Configure the maximum power for the PoE interface. | poe max-power <i>max-power</i> | Optional.
By default, 30000 milliwatts is the maximum power for the PoE interface for PoE+ switches. |

Configuring PoE power management

PoE power management involves PSE power management and PoE interface power management.

Configuring PoE interface power management

The power supply priority of a PD depends on the priority of the PoE interface. The priority levels of PoE interfaces are critical, high and low in descending order. Power supply to a PD is subject to PoE interface power management policies.

All PSEs implement the same PoE interface power management policies. When a PSE supplies power to a PD, the following actions occur:

- If the PoE interface power management is not enabled, no power will be supplied to a new PD when the PSE power is overloaded.
- If the PoE interface power management priority policy is enabled, the PD with a lower priority is first powered off to guarantee the power supply to the PD with a higher priority when the PSE power is overloaded.

19 watts guard band is reserved for each PoE interface on the device to prevent a PD from being powered off because of a sudden increase of the PD power. When the remaining power of the PSE where the PoE interface resides is lower than 19 watts and no priority is configured for the PoE interface, the PSE does not supply power to the new PD; when the remaining power of the PSE where the PoE interface resides is lower than 19 watts, but priority is configured for the PoE interface, the interface with a higher priority can preempt the power of the interface with a lower priority to ensure the normal working of the higher priority interface.

If the sudden increase of the PD power results in PSE power overload, power supply to the PD on the PoE interface with a lower priority will be stopped to ensure the power supply to the PD with a higher priority.

If the guaranteed remaining PSE power (the maximum PSE power minus the power allocated to the critical PoE interface, regardless of whether PoE is enabled for the PoE interface) is lower than the maximum power of the PoE interface, you will fail to set the priority of the PoE interface to **critical**. Otherwise, you can succeed in setting the priority to **critical**, and this PoE interface will preempt the power of other PoE interfaces with a lower priority level. In the latter case, the PoE interfaces whose power is preempted will be powered off, but their configurations will remain unchanged. When you change the priority of a PoE interface from critical to a lower level, the PDs connecting to other PoE interfaces will have an opportunity of being powered.

Configuration prerequisites

Enable PoE for PoE interfaces.

Configuration procedure

To configure PoE interface power management:

| Step | Command | Remarks |
|--|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Configure PoE interface power management priority policy. | poe pd-policy priority | By default, this policy is not configured. |
| 3. Enter PoE interface view. | interface <i>interface-type</i>
<i>interface-number</i> | N/A |
| 4. Configure the power supply priority for a PoE interface. | poe priority { critical high low } | Optional.
By default, low is the power supply priority for the PSE. |

Configuring the PoE monitoring function

With the PoE monitoring function enabled, the system monitors the parameter values related to PoE power supply, PSE, PD, and device temperature in real time. When a specific value exceeds the limited range, the system automatically takes some measures to protect itself.

Configuring PSE power monitoring

When the PSE power exceeds or drops below the specified threshold, the system sends trap messages.

To configure a power alarm threshold for the PSE:

| Step | Command | Remarks |
|---|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Configure a power alarm threshold for the PSE. | poe utilization-threshold
<i>utilization-threshold-value</i> pse
<i>pse-id</i> | Optional.
The default setting is 80%. |

Monitoring PD

When a PSE starts or ends power supply to a PD, the system sends a trap message.

Configuring PoE interface through PoE profile

You can configure a PoE interface either at the CLI or by using a PoE profile and applying the PoE profile to the specified PoE interface(s).

To configure a single PoE interface, configure it at the CLI; to configure PoE interfaces in batches, use a PoE profile.

A PoE profile is a collection of configurations that contain multiple PoE features. On large-scale networks, you can apply a PoE profile to multiple PoE interfaces, and these interfaces have the same PoE features. If the PoE interface connecting to a PD changes to another one, apply the PoE profile applied on the

originally connected interface to the currently connected interface instead of reconfiguring the features defined in the PoE profile one by one, simplifying the PoE configurations.

The device supports multiple PoE profiles. You can define PoE configurations based on each PD, save the configurations for different PDs into different PoE profiles, and apply the PoE profiles to the access interfaces of PDs accordingly.

Configuration guidelines

- If a PoE profile is applied, it cannot be deleted or modified before you cancel its application.
- The **poe max-power** *max-power* and **poe priority** { **critical** | **high** | **low** } commands must be configured in only one way, that is, either at the CLI or by configuring PoE profile.
- A PoE parameter on a PoE interface must be configured, modified and deleted in only one way. If a parameter configured in a way (for example, at the CLI) is then configured in the other way (for example, through PoE profile), the latter configuration fails and the original one is still effective. To make the latter configuration effective, you must cancel the original one first.

Configuring PoE profile

| Step | Command | Remarks |
|---|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Create a PoE profile, and enter PoE profile view. | poe-profile <i>profile-name</i> [<i>index</i>] | N/A |
| 3. Enable PoE for the PoE interface. | poe enable | By default, this function is disabled. |
| 4. Configure the maximum power for the PoE interface. | poe max-power <i>max-power</i> | Optional.
By default, 30000 milliwatts is the maximum power for the PoE interface for PoE+ switches. |
| 5. Configure power supply priority for the PoE interface. | poe priority { critical high low } | Optional.
By default, low is the power supply priority for the PoE interface. |

Applying PoE profile

You can apply a PoE profile in either system view or interface view. If you perform application to a PoE interface in both views, the latter application takes effect. To apply a PoE profile to multiple PoE interfaces, the system view is more efficient.

To apply the PoE profile in system view:

| Step | Command |
|---|---|
| 1. Enter system view. | system-view |
| 2. Apply the PoE profile to one or multiple PoE interfaces. | apply poe-profile { index <i>index</i> name <i>profile-name</i> }
interface <i>interface-range</i> |

To apply the PoE profile in interface view:

| Step | Command |
|--|--|
| 1. Enter system view. | system-view |
| 2. Enter PoE interface view. | interface <i>interface-type interface-number</i> |
| 3. Apply the PoE profile to the current PoE interface. | apply poe-profile { index <i>index</i> name <i>profile-name</i> } |

A PoE profile can be applied to multiple PoE interfaces, while a PoE interface can be applied with only one PoE profile.

Upgrading PSE processing software in service

You can upgrade the PSE processing software in service in either of the following two modes:

- **refresh mode**—Enables you to update the PSE processing software without deleting it. Normally, you can upgrade the PSE processing software in the refresh mode through the command line.
- **full mode**—Deletes the PSE processing software and reloads it. If the PSE processing software is damaged (in this case, you can execute none of PoE commands successfully), you can upgrade the PSE processing software in full mode to restore the PSE function.

An in-service PSE processing software upgrade may be unexpectedly interrupted (for example, an error results in device reboot). If you fail to upgrade the PSE processing software in full mode after reboot, you can power off the device and restart it before upgrading it in full mode again. After upgrade, restart the device manually to make the new PSE processing software take effect.

Configuration guidelines

To upgrade the PSE processing software in service:

| Step | Command |
|--|---|
| 1. Enter system view. | system-view |
| 2. Upgrade the PSE processing software in service. | poe update { full refresh } <i>filename</i> pse <i>pse-id</i> |

Displaying and maintaining PoE

| Task | Command | Remarks |
|---|--|-----------------------|
| Display PSE information. | display poe device [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display the power supplying state of the specified PoE interface. | display poe interface [<i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display power information for PoE interfaces. | display poe interface power [<i>interface-type interface-number</i>] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display information about PSE. | display poe pse [<i>pse-id</i>] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |

| Task | Command | Remarks |
|--|---|-----------------------|
| Display the power supply states of all PoE interfaces connected to the PSE. | display poe pse <i>pse-id</i> interface [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display power information for all PoE interfaces connected to the PSE. | display poe pse <i>pse-id</i> interface power [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display the configurations and applications of the PoE profile. | display poe-profile [index <i>index</i> name <i>profile-name</i>] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display the configurations and applications of the PoE profile applied to the specified PoE interface. | display poe-profile interface <i>interface-type</i> <i>interface-number</i> [{ begin exclude include } <i>regular-expression</i>] | Available in any view |

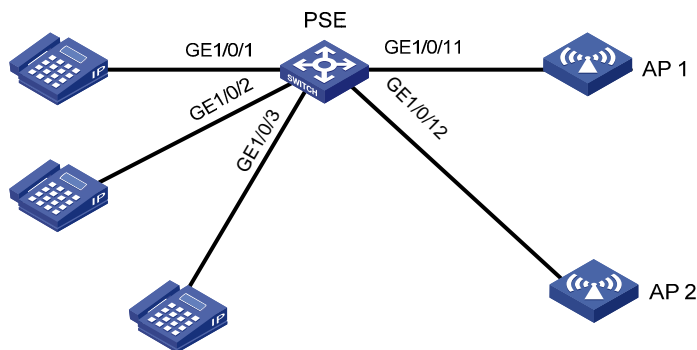
PoE configuration example

Network requirements

As shown in [Figure 56](#), the device supplies power to PDs through its PoE interfaces:

- GigabitEthernet 1/0/1, GigabitEthernet 1/0/2, and GigabitEthernet 1/0/3 are connected to IP telephones.
- GigabitEthernet 1/0/11 and GigabitEthernet 1/0/12 are connected to APs.
- The power supply priority of IP telephones is higher than that of the APs, for which the PSE supplies power to IP telephones first when the PSE power is overloaded.
- The maximum power of AP2 connected to GigabitEthernet 1/0/12 does not exceed 9000 milliwatts.

Figure 56 Network diagram



Configuration procedure

Enable PoE and specify the **critical** power supply priority on GigabitEthernet 1/0/1, GigabitEthernet 1/0/2, and GigabitEthernet 1/0/3.

```

<Sysname> system-view
[Sysname] interface gigabitethernet 1/0/1
[Sysname-GigabitEthernet1/0/1] poe enable
[Sysname-GigabitEthernet1/0/1] poe priority critical
  
```

```
[Sysname-GigabitEthernet1/0/1] quit
[Sysname] interface gigabitethernet 1/0/2
[Sysname-GigabitEthernet1/0/2] poe enable
[Sysname-GigabitEthernet1/0/2] poe priority critical
[Sysname-GigabitEthernet1/0/2] quit
[Sysname] interface gigabitethernet 1/0/3
[Sysname-GigabitEthernet1/0/3] poe enable
[Sysname-GigabitEthernet1/0/3] poe priority critical
[Sysname-GigabitEthernet1/0/3] quit
```

Enable PoE on GigabitEthernet 1/0/11 and GigabitEthernet 1/0/12, and configure the maximum power of GigabitEthernet 1/0/12 as 9000 milliwatts.

```
[Sysname] interface gigabitethernet 1/0/11
[Sysname-GigabitEthernet1/0/11] poe enable
[Sysname-GigabitEthernet1/0/11] quit
[Sysname] interface gigabitethernet 1/0/12
[Sysname-GigabitEthernet1/0/12] poe enable
[Sysname-GigabitEthernet1/0/12] poe max-power 9000
```

After the configuration takes effect, the IP telephones and AP devices are powered and can work normally.

Troubleshooting PoE

Setting the priority of a PoE interface to critical fails

Analysis

- The guaranteed remaining power of the PSE is lower than the maximum power of the PoE interface.
- The priority of the PoE interface is already set.

Solution

- In the first case, you can solve the problem by increasing the maximum PSE power, or by reducing the maximum power of the PoE interface when the guaranteed remaining power of the PSE cannot be modified.
- In the second case, you should first remove the priority already configured.

Applying a PoE profile to a PoE interface fails

Analysis

- Some configurations in the PoE profile are already configured.
- Some configurations in the PoE profile do not meet the configuration requirements of the PoE interface.
- Another PoE profile is already applied to the PoE interface.

Solution

- In the first case, you can solve the problem by removing the original configurations of those configurations.
- In the second case, you need to modify some configurations in the PoE profile.

- In the third case, you need to remove the application of the undesired PoE profile to the PoE interface.

Configuring an AC input under-voltage threshold fails

Analysis

The AC input under-voltage threshold is greater than or equal to the AC input over-voltage threshold.

Solution

You can drop the AC input under-voltage threshold below the AC input over-voltage threshold.

Configuring cluster management

Overview

Cluster management is an effective way to manage large numbers of dispersed network switches in groups. Cluster management offers the following advantages:

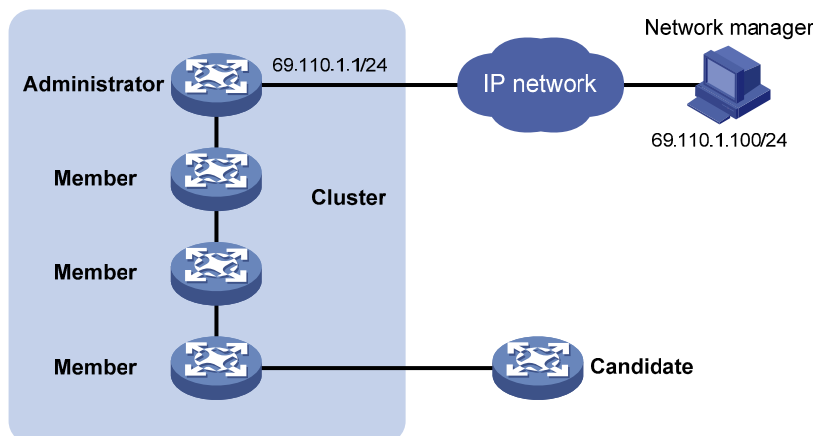
- Saves public IP address resources. You do not have to assign one public IP address for every cluster member switch.
- Simplifies configuration and management tasks. By configuring a public IP address on one switch, you can configure and manage a group of switches without the trouble of logging in to each switch separately.
- Provides a useful topology discovery and display function for network monitoring and debugging.
- Allows simultaneous software upgrading and parameter configuration on multiple switches, free from topology and distance limitations.

Roles in a cluster

The switches in a cluster play different roles according to their different functions and status. You can specify the following three roles for the switches:

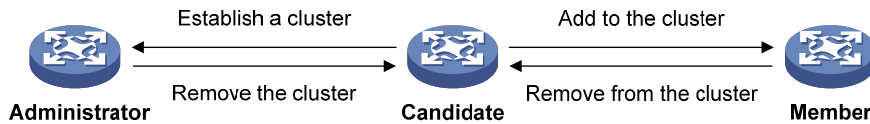
- **Management device (Administrator)**—A switch providing management interfaces for all switches in a cluster and the only switch configured with a public IP address. You can specify one and only one management switch for a cluster. Any configuration, management, and monitoring of the other switches in a cluster can only be implemented through the management switch. When a switch is specified as the management switch, it collects related information to discover and define candidate switches.
- **Member device (Member)**—A switch managed by the management switch in a cluster.
- **Candidate device (Candidate)**—A switch that does not yet belong to any cluster but can be added to a cluster. Different from a member switch, its topology information has been collected by the management switch but it has not been added to the cluster.

Figure 57 Network diagram



As shown in [Figure 57](#), the switch configured with a public IP address and performing the management function is the management switch, the other managed switches are member switches, and the switch that does not belong to any cluster but can be added to a cluster is a candidate switch. The management switch and the member switches form the cluster.

Figure 58 Role change in a cluster



As shown in [Figure 58](#), a switch in a cluster changes its role according to the following rules:

- A candidate switch becomes a management switch when you create a cluster on it. A management switch becomes a candidate switch only after the cluster is removed.
- A candidate switch becomes a member switch after being added to a cluster. A member switch becomes a candidate switch after it is removed from the cluster.

How a cluster works

Cluster management is implemented through HW Group Management Protocol version 2 (HGMPv2), which consists of the following three protocols:

- Neighbor Discovery Protocol (NDP)
- Neighbor Topology Discovery Protocol (NTDP)
- Cluster

A cluster configures and manages the switches in it through the above three protocols. Cluster management involves topology information collection and the establishment and maintenance of a cluster. Topology information collection and cluster maintenance are independent from each other; in fact, topology information collection starts before the cluster is created. The following workflow applies:

- All switches use NDP to collect the information of directly connected neighbors, including their software version, host name, MAC address and port number.
- The management switch uses NTDP to do the following:
 - Collect information about the switches within user-specified hops.
 - Collect the topology information of all switches.
 - Specifies the candidate switches of the cluster based on the collected information.
- The management switch adds or deletes a member switch and modifies cluster management configuration according to the candidate switch information collected through NTDP.

About NDP

NDP discovers information about directly connected neighbors, including the switch name, software version, and connecting port of the adjacent switches. NDP works in the following ways:

- A switch running NDP periodically sends NDP packets to its neighbors. An NDP packet carries NDP information (including the switch name, software version, and connecting port, etc.) and the holdtime, which is how long the receiving switches will keep the NDP information. At the same time, the switch also receives (but does not forward) NDP packets from its neighbors.
- A switch running NDP stores and maintains an NDP table. The switch creates an entry in the NDP table for each neighbor. If a new neighbor is found, meaning the switch receives an NDP packet sent by the neighbor for the first time, the switch adds an entry in the NDP table. If the NDP

information carried in the NDP packet is different from the stored information, the corresponding entry and holdtime in the NDP table are updated; otherwise, only the holdtime of the entry is updated. If an entry's holdtime expires (in other words, no NDP information from the neighbor is received to restart the hold time before it ages out), the entry is removed from the NDP table.

NDP runs on the data link layer, and supports different network layer protocols.

About NTDP

NTDP provides information required for cluster management; it collects topology information about the switches within the specified hop count. Based on the neighbor information stored in the neighbor table maintained by NDP, NTDP on the management switch advertises NTDP topology-collection requests to collect the NDP information of all the switches in a specific network range as well as the connection information of all its neighbors. The information collected will be used by the management switch or the network management software to implement required functions.

When a member switch detects a change on its neighbors through its NDP table, it informs the management switch through handshake packets. Then the management switch triggers its NTDP to collect specific topology information, so that its NTDP can discover topology changes promptly.

The management switch collects topology information periodically. You can also administratively launch a topology information collection. The process of topology information collection is as follows:

- The management switch periodically sends NTDP topology-collection request from the NTDP-enabled ports.
- Upon receiving the request, the switch sends NTDP topology-collection response to the management switch, copies this response packet on the NTDP-enabled port and sends it to the adjacent switch. Topology-collection response includes the basic information of the NDP-enabled switch and NDP information of all adjacent switches.
- The adjacent switch performs the same operation until the NTDP topology-collection request is sent to all the switches within specified hops.

To avoid concurrent responses to an NTDP topology-collection request causing congestion and deny of service on the management device, a delay mechanism was introduced. You configure the delay parameters for NTDP on the management device. As a result:

- Each device waits for a period of time before forwarding an NTDP topology-collection request on the first NTDP-enabled port.
- After the first NTDP-enabled port forwards the request, all other NTDP-enabled ports wait for a period of time and forward the NTDP topology-collection request.

Cluster management maintenance

1. Adding a candidate switch to a cluster

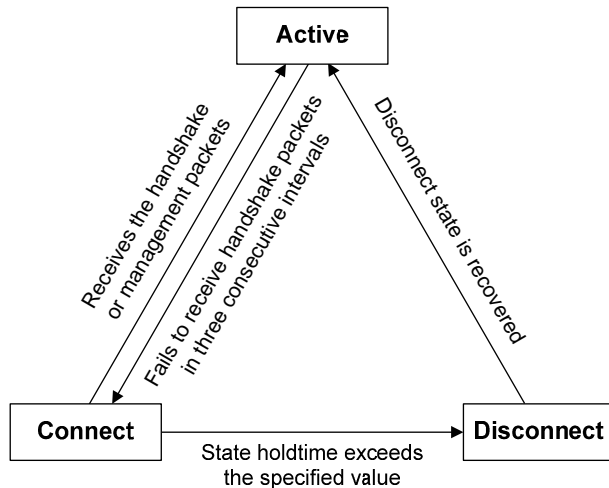
You should specify the management switch before creating a cluster. The management switch discovers and defines a candidate switch through NDP and NTDP protocols. The candidate switch can be automatically or manually added to the cluster.

After the candidate switch is added to the cluster, it can obtain the member number assigned by the management switch and the private IP address used for cluster management.

2. Communication within a cluster

In a cluster the management switch communicates with its member switches by sending handshake packets to maintain connection between them. The management/member switch state change is shown in [Figure 59](#).

Figure 59 Management/member switch state change



A cluster manages the state of its member devices as follows:

- After a cluster is created and a candidate switch is added to the cluster and becomes a member switch, the management switch saves the state information of the member switch and identifies it as Active. And the member switch also saves its state information and identifies itself as Active.
- After a cluster is created, its management switch and member switches begin to send handshake packets. Upon receiving the handshake packets from the other, the management switch or a member switch simply remains its state as Active, without sending a response.
- If the management switch does not receive the handshake packets from a member switch in an interval three times of the interval to send handshake packets, it changes the status of the member switch from Active to Connect. Likewise, if a member switch fails to receive the handshake packets from the management switch in an interval three times of the interval to send handshake packets, the status of itself will also be changed from Active to Connect.
- During information holdtime, if the management switch receives handshake or management packets from a member switch that is in Connect state, it changes the state of the member switch to Active. Otherwise, it considers the member switch to be disconnected, and changes the state of the member switch to Disconnect.
- During information holdtime, if a member switch in Connect state changes its state to Active if it receives handshake or management packets from the management switch; otherwise, it changes its state to Disconnect.
- If communication between the management switch and a member switch is recovered, the member switch which is in Disconnect state will be added to the cluster, and the state of the member switch locally and on the management switch will be changed to Active.
- Also, a member switch sends handshake packets to inform the management switch when there is a neighbor topology change.

Management VLAN

The management VLAN is a VLAN used for communication in a cluster; it limits the cluster management range. Through configuration of the management VLAN, the following functions can be implemented:

- Management packets (including NDP, NTDP and handshake packets) are restricted within the management VLAN. This isolates them from other packets, which enhances security.
- The management switch and the member switches communicate with each other through the management VLAN.

For a cluster to work normally, you must set the packets from the management VLAN to pass the ports connecting the management switch and the member/candidate switches (including the cascade ports). Therefore:

- If the packets from the management VLAN cannot pass a port, the switch connected with the port cannot be added to the cluster. Therefore, if the ports (including the cascade ports) connecting the management switch and the member/candidate switches prohibit the packets from the management VLAN, you can set the packets from the management VLAN to pass the ports on candidate switches with the management VLAN auto-negotiation function.
- Normally, only the packets with tags from the management VLAN can pass the ports. However, you can set packets without tags from the management VLAN to pass the ports if the default VLAN ID of the cascade ports and of the ports connecting the management switch and the member/candidate switches is the same as that of the management VLAN.

If a candidate switch is connected to a management switch through another candidate switch, the ports between the two candidate switches are cascade ports.

For more information about VLAN, see *Layer 2—LAN Switching Configuration Guide*.

Cluster management configuration task list

Before configuring a cluster, you need to determine the roles and functions the switches play. You also need to configure the related functions, preparing for the communication between switches within the cluster.

Configuration guidelines

- Disabling the NDP and NTDP functions on the management switch and member switches after a cluster is created will not cause the cluster to be dismissed, but will influence the normal operation of the cluster.
- In a cluster, if a member switch enabled with the 802.1X or MAC address authentication function has other member switches connected to it, you must enable HW Authentication Bypass Protocol (HABP) server on the switch. Otherwise, the management switch of the cluster cannot manage the switches connected with it. For more information about the HABP, see *Security Configuration Guide*.
- If the routing table of the management switch is full when a cluster is established, that is, entries with the destination address as a candidate switch cannot be added to the routing table, all candidate switches will be added to and removed from the cluster repeatedly.
- If the routing table of a candidate switch is full when the candidate switch is added to a cluster, that is, the entry with the destination address as the management switch cannot be added to the routing table, the candidate switch will be added to and removed from the cluster repeatedly.

Complete these tasks to configure cluster management functions:

| Task | Remarks |
|---|----------|
| Configuring the management switch: | |
| Enabling NDP globally and for specific ports | Optional |
| Configuring NDP parameters | Optional |
| Enabling NTDP globally and for specific ports | Optional |
| Configuring NTDP parameters | Optional |

| Task | Remarks |
|--|----------|
| Manually collecting topology information | Optional |
| Enabling the cluster function | Optional |
| Establishing a cluster | Required |
| Enabling management VLAN auto-negotiation | Required |
| Configuring communication between the management switch and the member switches within a cluster | Optional |
| Configuring cluster management protocol packets | Optional |
| Cluster member management | Optional |
| Configuring the member switches: | |
| Enabling NDP | Optional |
| Enabling NTDP | Optional |
| Manually collecting topology information | Optional |
| Enabling the cluster function | Optional |
| Deleting a member switch from a cluster | Optional |
| Configuring access between the management switch and its member switches | Optional |
| Adding a candidate switch to a cluster | Optional |
| Configuring advanced cluster management functions: | |
| Configuring topology management | Optional |
| Configuring interaction for a cluster | Optional |
| SNMP configuration synchronization function | Optional |
| Configuring web user accounts in batches | Optional |

Configuring the management switch

Perform the tasks in this section to configure the management switch for a cluster.

Enabling NDP globally and for specific ports

For NDP to work normally, you must enable NTDP both globally and on specific ports.

To enable NDP globally and for specific ports:

| Step | Command | Remarks |
|-------------------------|--------------------|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enable NDP globally. | ndp enable | Optional.
By default, this function is enabled. |

| Step | Command | Remarks |
|-------------------------------------|--|---|
| 3. Enable the NDP feature on ports. | <ul style="list-style-type: none"> • In system view:
ndp enable
interface <i>interface-list</i> • In Ethernet interface view or Layer 2 aggregate interface view: <ul style="list-style-type: none"> a. interface <i>interface-type</i>
<i>interface-number</i> b. ndp enable | Use either command.
By default, NDP is enabled globally and also on all ports. |

HP recommends that you disable NDP on a port which connects with the switches that do not need to join the cluster. This prevents the management switch from adding and collecting topology information from switches which do not need to join the cluster.

Configuring NDP parameters

A port enabled with NDP periodically sends NDP packets to its neighbors. If no NDP information from the neighbor is received to reset the holdtime, the holdtime times out and the switch removes the corresponding entry from the NDP table.

The time for the receiving switch to hold NDP packets cannot be shorter than the interval for sending NDP packets. Otherwise, the NDP table may become instable.

To configure NDP parameters:

| Step | Command | Remarks |
|---|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Configure the interval for sending NDP packets. | ndp timer hello <i>hello-time</i> | Optional.
The default interval is 60 seconds. |
| 3. Configure the period for the receiving switch to keep the NDP packets. | ndp timer aging <i>aging-time</i> | Optional.
The default setting is 180 seconds. |

Enabling NTDP globally and for specific ports

For NTDP to work normally, you must enable NTDP both globally and on specific ports.

To enable NTDP globally and for specific ports:

| Step | Command | Remarks |
|---|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enable NTDP globally. | ntdp enable | Optional.
By default, NTDP is enabled globally. |
| 3. Enter Ethernet interface view or Layer 2 aggregate interface view. | interface <i>interface-type</i>
<i>interface-number</i> | N/A |

| Step | Command | Remarks |
|------------------------------|--------------------|--|
| 4. Enable NTDP for the port. | ntdp enable | Optional.
By default, NTDP is enabled on all ports. |

HP recommends that you disable NTDP on a port which connects with the switches that do not need to join the cluster. This prevents the management switch from adding and collecting topology information from switches which do not need to join the cluster.

Configuring NTDP parameters

By configuring the maximum hops for collecting topology information, you can get topology information of the switches in a specified range, thus avoiding unlimited topology collection.

After the interval for collecting topology information is configured, the switch collects the topology information at this interval.

To avoid network congestion caused by large amounts of topology responses received in short periods:

- Upon receiving an NTDP topology-collection request, a switch does not forward it immediately. Instead, it waits for a period of time and then forwards the NTDP topology-collection request on its first NTDP-enabled port.
- Except for its first port, each switch's NTDP-enabled ports wait for a period of time, and then forward the NTDP topology collection request after the previous port forwards it.

To configure NTDP parameters:

| Step | Command | Remarks |
|--|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Configure the maximum hops for topology collection. | ntdp hop <i>hop-value</i> | Optional.
The default setting is 3. |
| 3. Configure the interval for collecting topology information. | ntdp timer <i>interval</i> | Optional.
The default interval is 1 minute. |
| 4. Configure the delay to forward topology-collection request packets on the first port. | ntdp timer hop-delay <i>delay-time</i> | Optional.
The default setting is 200 ms. |
| 5. Configure the port delay to forward topology-collection request on other ports. | ntdp timer port-delay <i>delay-time</i> | Optional.
The default setting is 20 ms. |

The two delay values should be configured on the topology collecting switch. A topology-collection request sent by the topology collecting switch carries the two delay values, and a switch that receives the request forwards the request according to the delays.

Manually collecting topology information

The management switch collects topology information periodically after a cluster is created. In addition, you can manually start topology information collection on the management switch or NTDP-enabled

switch, thus managing and monitoring switches in real time, regardless of whether a cluster is created. To configure to manually collect topology information:

| Task | Command |
|--|---------------------|
| Manually collect topology information. | ntdp explore |

Enabling the cluster function

| Step | Command | Remarks |
|--|-----------------------|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enable the cluster function globally. | cluster enable | Optional.
By default, this function is enabled. |

Establishing a cluster

Before establishing a cluster, you need to specify the management VLAN, and you cannot modify the management VLAN after a device is added to the cluster.

In addition, you need to configure a private IP address pool for the devices to be added to the cluster on the device to be configured as the management device before establishing a cluster. Meanwhile, the IP addresses of the VLAN interfaces of the management device and member devices cannot be in the same network segment as that of the cluster address pool; otherwise, the cluster cannot work normally. When a candidate device is added to a cluster, the management device assigns it a private IP address for it to communicate with other devices in the cluster.

You can establish a cluster in two ways: manually and automatically. You can follow the prompts to establish a cluster automatically. The system cluster auto-establishment process will prompt you through the following steps:

1. Enter a name for the cluster you want to establish.
2. List all the candidate switches within your predefined hop count.
3. Start to automatically add them to the cluster.

You can press **Ctrl+C** anytime during the adding process to exit the cluster auto-establishment process. However, this will only stop adding new switches into the cluster, and switches already added into the cluster are not removed.

To establish a cluster manually or automatically:

| Step | Command | Remarks |
|--|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Specify the management VLAN. | management-vlan <i>vlan-id</i> | Optional.
By default, VLAN 1 is the management VLAN. |
| 3. Enter cluster view. | cluster | N/A |
| 4. Configure the private IP address range for member switches. | ip-pool <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } | By default, the private IP address range for member switches is not configured. |

| Step | Command | Remarks |
|-------------------------|---|---|
| 5. Establish a cluster. | <ul style="list-style-type: none"> Manually establish a cluster:
build <i>cluster-name</i> Automatically establish a cluster:
auto-build [recover] | <p>Use either approach.</p> <p>By default, the switch is not the management switch.</p> |

Handshake packets use UDP port 40000. For a cluster to be established successfully, make sure that the port is not in use before establishing it.

Enabling management VLAN auto-negotiation

The management VLAN limits the cluster management range. If the switch discovered by the management switch does not belong to the management VLAN, meaning the cascade ports and the ports connecting with the management switch do not allow the packets from the management VLAN to pass, and the new switch cannot be added to the cluster. Through the configuration of the management VLAN auto-negotiation function, the cascade ports and the ports directly connected to the management switch can be automatically added to the management VLAN.

Configuration guidelines

When the management VLAN auto-negotiation is enabled, the ports connecting member switches change as follows:

- If a port was an access port, after changing to a hybrid port, the port does not permit the packets of any other VLAN to pass through except the management VLAN, which passes as tagged.
- If a port was a trunk or a hybrid port, the link type change process does not affect the port type and the permitted VLANs. The only change is permit the packets of the management VLAN to pass through tagged only for a hybrid port.

Before enabling this function, check the link types of ports connecting member switches and the VLANs whose packets are permitted to pass through to avoid influence to your network due to link type change of ports.

To configure management VLAN auto-negotiation:

| Step | Command | Remarks |
|---|---|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter cluster view. | cluster | N/A |
| 3. Enable management VLAN auto-negotiation. | management-vlan synchronization
enable | By default, this function is disabled. |

Configuring communication between the management switch and the member switches within a cluster

In a cluster, the management switch and member switches communicate by sending handshake packets to maintain connection between them. You can configure interval of sending handshake packets and the holdtime of a switch on the management switch. This configuration applies to all member switches within the cluster. For a member switch in Connect state:

- If the management switch does not receive handshake packets from a member switch within the

holdtime, it changes the state of the member switch to Disconnect. When the communication is recovered, the member switch needs to be re-added to the cluster (this process is automatically performed).

- If the management switch receives handshake packets from the member switch within the holdtime, the state of the member switch remains Active and the holdtime is restarted.

To configure communication between the management switch and the member switches within a cluster:

| Step | Command | Remarks |
|--|----------------------------------|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter cluster view. | cluster | N/A |
| 3. Configure the handshake interval. | timer <i>interval</i> | Optional.
The default interval is 10 seconds. |
| 4. Configure the holdtime of a switch. | holdtime <i>hold-time</i> | Optional.
The default setting is 60 seconds. |

Configuring cluster management protocol packets

By default, the destination MAC address of cluster management protocol packets (including NDP, NTDP and HABP packets) is a multicast MAC address 0180-C200-000A, which IEEE reserved for later use. Since some switches cannot forward the multicast packets with the destination MAC address of 0180-C200-000A, so cluster management packets cannot traverse these switches. For a cluster to work normally in this case, you can modify the destination MAC address of a cluster management protocol packet without changing the current networking.

The management switch periodically sends MAC address negotiation broadcast packets to advertise the destination MAC address of the cluster management protocol packets.

When you configure the destination MAC address for cluster management protocol packets:

- If the interval for sending MAC address negotiation broadcast packets is 0, the system automatically sets it to 1 minute.
- If the interval for sending MAC address negotiation broadcast packets is not 0, the interval remains unchanged.

To configure the destination MAC address of the cluster management protocol packets:

| Step | Command | Remarks |
|------------------------|--------------------|---------|
| 1. Enter system view. | system-view | N/A |
| 2. Enter cluster view. | cluster | N/A |

| Step | Command | Remarks |
|---|---|---|
| 3. Configure the destination MAC address for cluster management protocol packets. | cluster-mac <i>mac-address</i> | By default, the destination MAC address is 0180-C200-000A.
The following are the configurable MAC addresses: <ul style="list-style-type: none"> • 0180-C200-0000 • 0180-C200-000A • 0180-C200-0020 through 0180-C200-002F • 010F-E200-0002 |
| 4. Configure the interval for sending MAC address negotiation broadcast packets. | cluster-mac syn-interval <i>interval</i> | Optional.
The default interval is one minute. |

Cluster member management

You can manually add a candidate switch to a cluster, or remove a member switch from a cluster.

If a member switch needs to be rebooted for software upgrade or configuration update, you can remotely reboot it through the management switch.

Adding a member switch

| Step | Command |
|---|--|
| 1. Enter system view. | system-view |
| 2. Enter cluster view. | cluster |
| 3. Add a candidate switch to the cluster. | add-member [<i>member-number</i>] mac-address <i>mac-address</i> [password <i>password</i>] |

Removing a member switch

| Step | Command |
|---|--|
| 1. Enter system view. | system-view |
| 2. Enter cluster view. | cluster |
| 3. Remove a member switch from the cluster. | delete-member <i>member-number</i> [to-black-list] |

Rebooting a member switch

| Step | Command |
|--------------------------------------|---|
| 1. Enter system view. | system-view |
| 2. Enter cluster view. | cluster |
| 3. Reboot a specified member switch. | reboot member { <i>member-number</i> mac-address <i>mac-address</i> } [eraseflash] |

Configuring the member switches

Enabling NDP

See "Enabling NDP globally and for specific ports."

Enabling NTDP

See "Enabling NTDP globally and for specific ports."

Manually collecting topology information

See "Manually collecting topology information."

Enabling the cluster function

See "Enabling the cluster function."

Deleting a member switch from a cluster

| Step | Command |
|---|-----------------------------------|
| 1. Enter system view. | system-view |
| 2. Enter cluster view. | cluster |
| 3. Delete a member switch from the cluster. | undo administrator-address |

Configuring access between the management switch and its member switches

After having successfully configured NDP, NTDP and cluster, you can configure, manage and monitor the member switches through the management switch. You can manage member switches in a cluster through switching from the operation interface of the management switch to that of a member switch or configure the management switch by switching from the operation interface of a member switch to that of the management switch.

Configuration guidelines

Telnet connection is used in the switching between the management switch and a member switch. Note the following when switching between them:

- Authentication is required when you switch from a member switch to the management switch. The switching fails if authentication is not passed. If authentication is passed, your user level is allocated by the management switch according to the predefined level.
- When a candidate switch is added to a cluster and becomes a member switch, its super password with the level of 3 will be automatically synchronized to the management switch. Therefore, after a cluster is established, it is not recommended to modify the super password of any member

(including the management switch and member switches) of the cluster; otherwise, the switching may fail because of an authentication failure.

- If the member specified in this command does not exist, the system prompts error when you execute the command; if the switching succeeds, your user level on the management switch is retained.
- If the Telnet users on the switch to be logged in reach the maximum number, the switching fails.
- To prevent resource waste, avoid ring switching when configuring access between cluster members. For example, if you switch from the operation interface of the management switch to that of a member switch and then need to switch back to that of the management switch, use the **quit** command to end the switching, but not the **cluster switch-to administrator** command to switch to the operation interface of the management switch.

Configuration procedure

To configure access between member switches of a cluster:

| Step | Command |
|---|--|
| 1. Switch from the operation interface of the management switch to that of a member switch. | cluster switch-to { <i>member-number</i> mac-address <i>mac-address</i> sysname <i>member-sysname</i> } |
| 2. Switch from the operation interface of a member switch to that of the management switch. | cluster switch-to administrator |

Adding a candidate switch to a cluster

| Step | Command |
|---|---|
| 1. Enter system view. | system-view |
| 2. Enter cluster view. | cluster |
| 3. Add a candidate switch to the cluster. | administrator-address <i>mac-address</i> name <i>name</i> |

Configuring advanced cluster management functions

Configuring topology management

The concepts of blacklist and whitelist are used for topology management. An administrator can diagnose the network by comparing the current topology (namely, the information of a node and its neighbors in the cluster) and the standard topology.

- Topology management whitelist (standard topology): A whitelist is a list of topology information that has been confirmed by the administrator as correct. You can get the information of a node and its neighbors from the current topology. Based on the information, you can manage and maintain the whitelist by adding, deleting or modifying a node.
- Topology management blacklist: Switches in a blacklist are not allowed to join a cluster. A blacklist contains the MAC addresses of switches. If a blacklisted switch is connected to a network through

another switch not included in the blacklist, the MAC address and access port of the latter are also included in the blacklist. The candidate switches in a blacklist can be added to a cluster only if the administrator manually removes them from the list.

The whitelist and blacklist are mutually exclusive. A whitelist member cannot be a blacklist member, and the blacklist member cannot be a whitelist member. However, a topology node can belong to neither the whitelist nor the blacklist. Nodes of this type are usually newly added nodes, whose identities are to be confirmed by the administrator.

You can back up and restore the whitelist and blacklist in the following two ways:

- Backing them up on the FTP server shared by the cluster. You can manually restore the whitelist and blacklist from the FTP server.
- Backing them up in the Flash of the management switch. When the management switch restarts, the whitelist and blacklist will be automatically restored from the Flash. When a cluster is re-established, you can choose whether to restore the whitelist and blacklist from the Flash automatically, or you can manually restore them from the Flash of the management switch.

To configure cluster topology management:

| Step | Command | Remarks |
|---|---|-----------|
| 1. Enter system view. | system-view | N/A |
| 2. Enter cluster view. | cluster | N/A |
| 3. Add a switch to the blacklist. | black-list add-mac <i>mac-address</i> | Optional. |
| 4. Remove a switch from the blacklist. | black-list delete-mac { all <i>mac-address</i> } | Optional. |
| 5. Confirm the current topology and save it as the standard topology. | topology accept { all [save-to { ftp-server local-flash }] mac-address <i>mac-address</i> member-id <i>member-number</i> } | Optional. |
| 6. Save the standard topology to the FTP server or the local Flash. | topology save-to { ftp-server local-flash } | Optional. |
| 7. Restore the standard topology information. | topology restore-from { ftp-server local-flash } | Optional. |

Configuring interaction for a cluster

After establishing a cluster, you can configure FTP/TFTP server, NM host and log host for the cluster on the management switch.

- After you configure an FTP/TFTP server for a cluster, the members in the cluster access the FTP/TFTP server configured through the management switch. Execute the **ftp server-address** or **tftp server-address** command and specifying the private IP address of the management switch as the *server-address*. For more information about the **ftp** and **tftp** commands, see *Fundamentals Command Reference*.
- After you configure a log host for a cluster, all the log information of the members in the cluster will be output to the configured log host in the following way:
 - Member switches send their log information to the management switch.
 - The management switch converts the addresses of log information and sends them to the log host.

- After you configure an NM host for a cluster, the member switches in the cluster send their Trap messages to the shared SNMP NM host through the management switch.

If the port of an access NM switch (including FTP/TFTP server, NM host and log host) does not allow the packets from the management VLAN to pass, the NM switch cannot manage the switches in a cluster through the management switch. In this case, on the management switch, you need to configure the VLAN interface of the access NM switch (including FTP/TFTP server, NM host and log host) as the NM interface.

To isolate management protocol packets of a cluster from packets outside the cluster, HP recommends you configure the ports connected to the external networks as not allowing the management VLAN to pass through and configure the NM interface for the management switch.

To configure the interaction for a cluster:

| Step | Command | Remarks |
|---|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter cluster view. | cluster | N/A |
| 3. Configure the FTP server shared by the cluster. | ftp-server <i>ip-address</i> [user-name <i>username</i> password { simple cipher } <i>password</i>] | By default, no FTP server is configured for a cluster. |
| 4. Configure the TFTP server shared by the cluster. | tftp-server <i>ip-address</i> | By default, no TFTP server is configured for a cluster. |
| 5. Configure the log host shared by the member switches in the cluster. | logging-host <i>ip-address</i> | By default, no log host is configured for a cluster. |
| 6. Configure the SNMP NM host shared by the cluster. | snmp-host <i>ip-address</i> [community-string read <i>string1</i> write <i>string2</i>] | By default, no SNMP host is configured. |
| 7. Configure the NM interface of the management switch. | nm-interface vlan-interface <i>interface-name</i> | Optional. |

SNMP configuration synchronization function

Using the SNMP configuration synchronization function facilitates management of a cluster, with which you can perform SNMP-related configurations on the management switch and synchronize them to the member switches on the whitelist. This operation is equal to configuring multiple member switches at one time. It simplifies the configuration process.

To configure the SNMP configuration synchronization function:

| Step | Command | Remarks |
|---|--|---------|
| 1. Enter system view. | system-view | N/A |
| 2. Enter cluster view. | cluster | N/A |
| 3. Configure the SNMP community name shared by a cluster. | cluster-snmp-agent community { read write } <i>community-name</i> [mib-view <i>view-name</i>] | N/A |

| Step | Command | Remarks |
|--|--|---|
| 4. Configure the SNMPv3 group shared by a cluster. | cluster-snmp-agent group v3
<i>group-name</i> [authentication privacy] [read-view <i>read-view</i>] [write-view <i>write-view</i>] [notify-view <i>notify-view</i>] | N/A |
| 5. Create or update information of the MIB view shared by a cluster. | cluster-snmp-agent mib-view included <i>view-name oid-tree</i> | By default, the name of the MIB view shared by a cluster is ViewDefault and a cluster can access the ISO subtree. |
| 6. Add a user for the SNMPv3 group shared by a cluster. | cluster-snmp-agent usm-user v3
<i>user-name group-name</i>
[authentication-mode { md5 sha } <i>auth-password</i>] [privacy-mode des56 <i>priv-password</i>] | N/A |

The SNMP-related configurations are retained when a cluster is dismissed or the member switches are removed from the whitelist. For more information about SNMP, see "[Configuring SNMP](#)."

Configuring web user accounts in batches

Configuring web user accounts in batches enables you to do the following:

- Through the web interface, configure, on the management switch, the username and password used to log in to the cluster switches (including the management switch and member switches).
- Synchronize the configurations to the member switches in the whitelist.

This operation is equal to performing the configurations on the member switches. You need to enter your username and password when you log in to the cluster switches (including the management switch and member switches) through the web interface.

To configure web user accounts in batches:

| Step | Command |
|--|---|
| 1. Enter system view. | system-view |
| 2. Enter cluster view. | cluster |
| 3. Configure web user accounts in batches. | cluster-local-user <i>user-name password</i> { cipher simple } <i>password</i> |

If a cluster is dismissed or the member switches are removed from the whitelist, the configurations of web user accounts are still retained.

Displaying and maintaining cluster management

| Task | Command | Remarks |
|--|--|-----------------------|
| Display NDP configuration information. | display ndp [interface <i>interface-list</i>] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |

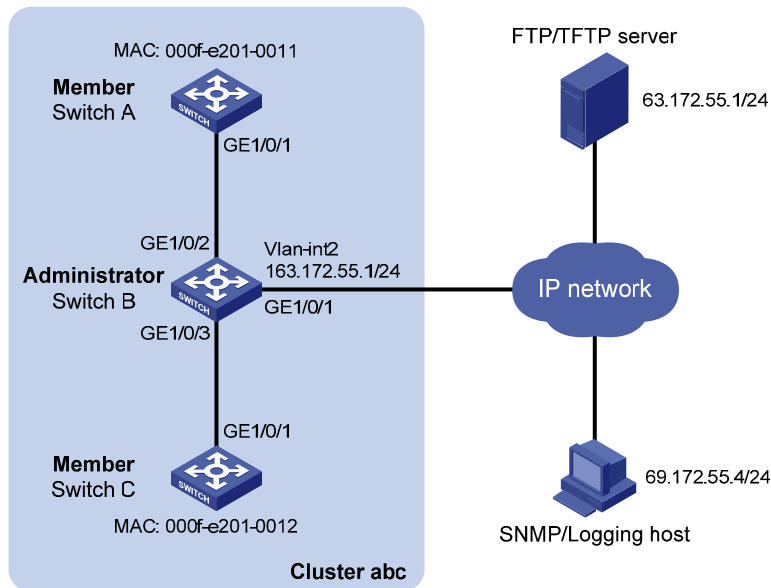
| Task | Command | Remarks |
|---|---|------------------------|
| Display NTDP configuration information. | display ntdp [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display the switch information collected through NTDP. | display ntdp device-list [verbose] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display the detailed NTDP information of a specified switch. | display ntdp single-device mac-address <i>mac-address</i> [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display information of the cluster to which the current switch belongs. | display cluster [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display the standard topology information. | display cluster base-topology [<i>mac-address mac-address</i> member-id <i>member-number</i>] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display the current blacklist of the cluster. | display cluster black-list [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display the information of candidate switches. | display cluster candidates [<i>mac-address mac-address</i> verbose] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display the current topology information. | display cluster current-topology [<i>mac-address mac-address</i> [to-mac-address <i>mac-address</i>] member-id <i>member-number</i> [to-member-id <i>member-number</i>]] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display information about cluster members. | display cluster members [<i>member-number</i> verbose] [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Clear NDP statistics. | reset ndp statistics [interface <i>interface-list</i>] | Available in user view |

Cluster management configuration example

Network requirements

- Three switches form cluster **abc**, whose management VLAN is VLAN 10. In the cluster, Switch B serves as the management switch (Administrator), whose network management interface is VLAN-interface 2; Switch A and Switch C are the member switches (Member).
- All the switches in the cluster use the same FTP server and TFTP server on host 63.172.55.1/24, and use the same SNMP NMS and log services on host IP address: 69.172.55.4/24.
- Add the switch whose MAC address is 000f-e201-0013 to the blacklist.

Figure 60 Network diagram



Configuration procedure

1. Configure the member switch Switch A:

Enable NDP globally and for port GigabitEthernet 1/0/1.

```
<SwitchA> system-view
[SwitchA] ndp enable
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] ndp enable
[SwitchA-GigabitEthernet1/0/1] quit
```

Enable NTPD globally and for port GigabitEthernet 1/0/1.

```
[SwitchA] ntpd enable
[SwitchA] interface gigabitethernet 1/0/1
[SwitchA-GigabitEthernet1/0/1] ntpd enable
[SwitchA-GigabitEthernet1/0/1] quit
```

Enable the cluster function.

```
[SwitchA] cluster enable
```

2. Configure the member switch Switch C:

As the configurations for the member switches are the same, the configuration procedure for Switch C is not shown here.

3. Configure the management switch Switch B:

Enable NDP globally and for ports GigabitEthernet 1/0/2 and GigabitEthernet 1/0/3.

```
<SwitchB> system-view
[SwitchB] ndp enable
[SwitchB] interface gigabitethernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] ndp enable
[SwitchB-GigabitEthernet1/0/2] quit
[SwitchB] interface gigabitethernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] ndp enable
```

```

[SwitchB-GigabitEthernet1/0/3] quit
# Configure the period for the receiving switch to keep NDP packets as 200 seconds.
[SwitchB] ndp timer aging 200
# Configure the interval to send NDP packets as 70 seconds.
[SwitchB] ndp timer hello 70
# Enable NTDP globally and for ports GigabitEthernet 1/0/2 and GigabitEthernet 1/0/3.
[SwitchB] ntdp enable
[SwitchB] interface gigabitethernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] ntdp enable
[SwitchB-GigabitEthernet1/0/2] quit
[SwitchB] interface gigabitethernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] ntdp enable
[SwitchB-GigabitEthernet1/0/3] quit
# Configure the hop count to collect topology as 2.
[SwitchB] ntdp hop 2
# Configure the delay to forward topology-collection request packets on the first port as 150 ms.
[SwitchB] ntdp timer hop-delay 150
# Configure the delay to forward topology-collection request packets on the first port as 15 ms.
[SwitchB] ntdp timer port-delay 15
# Configure the interval to collect topology information as 3 minutes.
[SwitchB] ntdp timer 3
# Configure the management VLAN of the cluster as VLAN 10.
[SwitchB] vlan 10
[SwitchB-vlan10] quit
[SwitchB] management-vlan 10
# Configure ports GigabitEthernet 1/0/2 and GigabitEthernet 1/0/3 as Trunk ports and allow
# packets from the management VLAN to pass.
[SwitchB] interface gigabitethernet 1/0/2
[SwitchB-GigabitEthernet1/0/2] port link-type trunk
[SwitchB-GigabitEthernet1/0/2] port trunk permit vlan 10
[SwitchB-GigabitEthernet1/0/2] quit
[SwitchB] interface gigabitethernet 1/0/3
[SwitchB-GigabitEthernet1/0/3] port link-type trunk
[SwitchB-GigabitEthernet1/0/3] port trunk permit vlan 10
[SwitchB-GigabitEthernet1/0/3] quit
# Enable the cluster function.
[SwitchB] cluster enable
# Configure a private IP address range for the member switches, which is from 172.16.0.1 to
# 172.16.0.7.
[SwitchB] cluster
[SwitchB-cluster] ip-pool 172.16.0.1 255.255.255.248
# Configure the current switch as the management switch, and establish a cluster named abc.
[SwitchB-cluster] build abc
Restore topology from local flash file,for there is no base topology.
(Please confirm in 30 seconds, default No). (Y/N)
N

```

```

# Enable management VLAN auto-negotiation.
[abc_0.SwitchB-cluster] management-vlan synchronization enable
# Configure the holdtime of the member switch information as 100 seconds.
[abc_0.SwitchB-cluster] holdtime 100
# Configure the interval to send handshake packets as 10 seconds.
[abc_0.SwitchB-cluster] timer 10
# Configure the FTP Server, TFTP Server, Log host and SNMP host for the cluster.
[abc_0.SwitchB-cluster] ftp-server 63.172.55.1
[abc_0.SwitchB-cluster] tftp-server 63.172.55.1
[abc_0.SwitchB-cluster] logging-host 69.172.55.4
[abc_0.SwitchB-cluster] snmp-host 69.172.55.4
# Add the switch whose MAC address is 000f-e201-0013 to the blacklist.
[abc_0.SwitchB-cluster] black-list add-mac 000f-e201-0013
[abc_0.SwitchB-cluster] quit
# Add port GigabitEthernet 1/0/1 to VLAN 2, and configure the IP address of VLAN-interface 2.
[abc_0.SwitchB] vlan 2
[abc_0.SwitchB-vlan2] port gigabitethernet 1/0/1
[abc_0.SwitchB] quit
[abc_0.SwitchB] interface vlan-interface 2
[abc_0.SwitchB-Vlan-interface2] ip address 163.172.55.1 24
[abc_0.SwitchB-Vlan-interface2] quit
# Configure VLAN-interface 2 as the network management interface.
[abc_0.SwitchB] cluster
[abc_0.SwitchB-cluster] nm-interface vlan-interface 2

```

Configuring CWMP

Hardware compatibility

The CWMP feature is available only on the HP 5500 EI switches.

CWMP overview

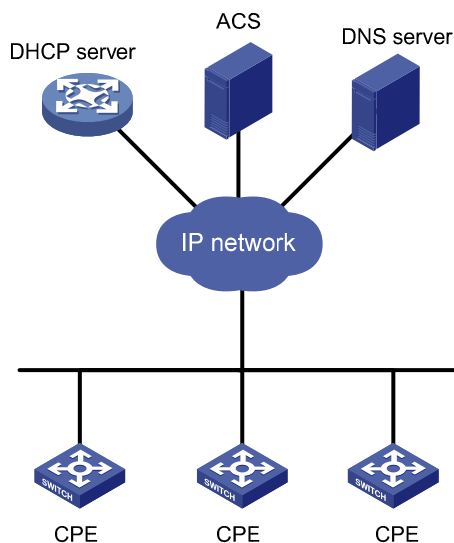
CPE WAN Management Protocol (CWMP), also called "TR-069," is a DSL Forum technical specification for remote management of home network devices. It defines the general framework, message format, management method, and data model for managing and configuring home network devices.

CWMP is mainly applied to DSL access networks, which are hard to manage because end-user devices are dispersed and large in number. CWMP makes the management easier by using an auto-configuration server (ACS) to perform remote centralized management of customer premises equipment (CPE).

CWMP network framework

Figure 61 shows a basic CWMP network framework. The DHCP server, ACS, and DNS server can also be located on the same server.

Figure 61 CWMP network framework



The basic CWMP network elements include:

- **CPE**—Customer premises equipment, the managed device in the network.
- **ACS**—Auto-configuration server, the management device in the network. In this document, ACS refers to the server installed with the HP iMC BIMS system.
- **DNS server**—Domain name system server. An ACS and a CPE use URLs to identify and access each other. DNS is used to resolve the URLs.

- **DHCP server**—Dynamic Host Configuration Protocol (DHCP) server, which assigns IP addresses to CPEs, and uses the options filed in the DHCP packet to provide configuration parameters (such as URL) to the CPEs.

Basic CWMP functions

Automatic configuration file deployment

The network administrator can create different configuration files on the ACS for access switches according to their service functions to realize fast configuration. After a connection is established between the ACS and a CPE, the ACS determines the type of the CPE and delivers the corresponding configuration file to the CPE. In this way, CPEs of the same type obtain the same service configurations. The ACS divides CPEs by their switch models or serial IDs.

A configuration file delivered by the ACS can be either the startup configuration or the running configuration on the CPE.

- **Startup configuration**—The configuration file delivered by the ACS overwrites the default configuration file on the CPE. After the CPE reboots, it runs the new configuration file.
- **Running configuration**—The configuration file delivered by the ACS is written to the running configuration file on the CPE, and the new configurations take effect immediately. You need to save the new configurations to make them survive a switch reboot.

CPE system software image file and configuration file management

The network administrator can save important files such as the application file and configuration file of a CPE to an ACS. If the ACS finds that a file is updated, it notifies the CPE to download the file by sending a request. After the CPE receives the request, it automatically downloads the file from the specified file server according to the filename and downloading address provided in the ACS request. After the CPE downloads the file, it checks the file validity and then report the download result (success or failure) to the ACS.

CPEs can download the following types of files from the ACS: application file and configuration file.

To backup important data, a CPE can upload the current configuration file and log files to the specified server according to the requirement of an ACS.

CPE status and performance monitoring

An ACS can monitor the parameters of a CPE connected to it. Different CPEs have different performances and functionalities. Therefore the ACS must be able to identify each type of CPE and monitor the current configuration and configuration changes of each CPE. CWMP also allows the administrator to define monitor parameters and get the parameter values through an ACS, so as to get the CPE status and statistics.

The status and performance that can be monitored by an ACS include:

- Manufacture name (Manufacturer)
- ManufacturerOUI
- SerialNumber
- HardwareVersion
- SoftwareVersion
- DeviceStatus
- UpTime
- Configuration file (ConfigFile)

- ACS address (URL)
- ACS username (Username)
- ACS password (Password)
- PeriodicInformEnable
- PeriodicInformInterval
- PeriodicInformTime
- CPE address (ConnectionRequestURL)
- CPE username (ConnectionRequestUsername)
- CPE password (ConnectionRequestPassword)

CWMP mechanism

Auto-connection between the ACS and a CPE

When a CPE starts up for the first time, it automatically obtains an IP address from the DHCP server, which informs the CPE of the following information:

- The URL address of the ACS (assigned by the DHCP server through Option 43)
- Username and password for connecting the ACS (assigned by the DHCP server through Option 43)
- DNS server address (directly assigned)

After the CPE receives the above information, it has the IP address of the ACS resolved by the DNS server, and sends a connection request to the ACS. If the CPE passes the authentication with the acquired username and password, the connection between the ACS and the CPE is established.

If the current session is not finished but the connection between ACS and CPE is interrupted, the CPE automatically establishes a new connection with the ACS until the number of CPE auto-connection retries reaches the limit.

The CPE can send connection requests either periodically or at the specified time to the ACS. The ACS can initiate a connection request to the CPE at any time, and can establish a connection with the CPE after passing CPE authentication.

Configuration parameter deployment

When a CPE logs in to an ACS, the ACS can automatically apply some configurations to the CPE for it to perform auto configuration. [Table 9](#) lists the auto-configuration parameters supported by the switch.

Table 9 Auto-configuration parameters and their functions

| Auto-configuration parameters | Function |
|---------------------------------|--|
| Configuration file (ConfigFile) | Updates the local configuration file on the CPE. The ACS delivers a configuration file to the CPE in one of the following formats: file or current configuration. |
| ACS address (URL) | Updates the ACS address kept on the CPE. The parameter is used when there is an active and standby ACS switchover. |
| ACS username (Username) | Automatically synchronizes the username and password on the CPE when those on the ACS change. The parameters are also used to inform the CPE of the authentication information of the standby ACS server when there is an active and standby ACS switchover. |
| ACS password (Password) | |
| PeriodicInformEnable | Enables the sending of Inform messages. |

| Auto-configuration parameters | Function |
|---|--|
| PeriodicInformInterval | Configures the CPE to send an Inform message periodically. The parameter is used for querying updates and information backup regularly. |
| PeriodicInformTime | Configures the CPE to send an Inform message at a specified time. The parameter is used for querying updates and information backup at a specified time. |
| CPE username
(ConnectionRequestUsername) | Configures the CPE username and password for connection to the ACS. |
| CPE password
(ConnectionRequestPassword) | |

RPC methods

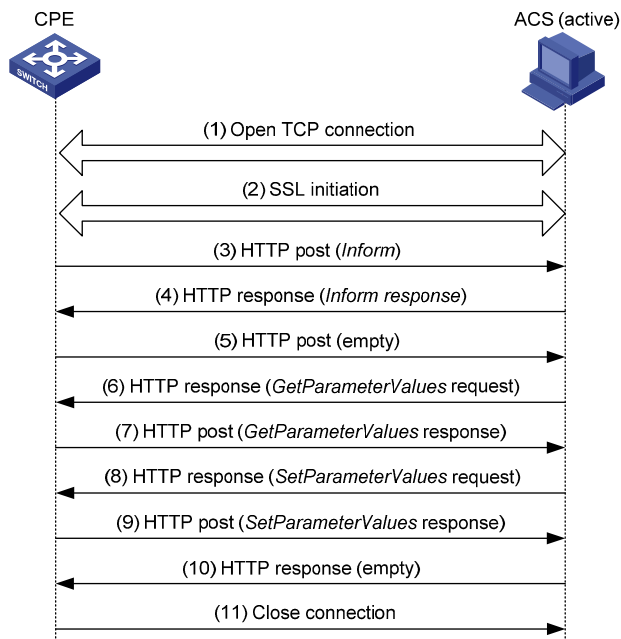
CWMP provides the following major remote procedure call (RPC) methods for an ACS to manage or monitor a CPE:

- **Get**—The ACS gets the value of one or more parameters from the CPE.
- **Set**—The ACS sets the value of one or more parameters on the CPE.
- **Inform**—The CPE sends an Inform message to an ACS whenever the CPE initiates a connection to the ACS, or the CPE's underlying configuration changes, or the CPE periodically sends its local information to the ACS.
- **Download**—The ACS requires a CPE to download a specific file from the specified URL, ensuring upgrading of CPE software and auto download of the vendor configuration file.
- **Upload**—The ACS requires a CPE to upload a specific file to the specified location.
- **Reboot**—The ACS remotely reboots the CPE when the CPE encounters a failure or completes a software upgrade.

How CWMP works

The following example illustrates how CWMP works. This example has two ACSs, main and backup in an area. The main ACS must restart for system upgrade. To ensure continuous CPE monitoring, the main ACS switch all CPEs in the area to the backup ACS.

Figure 62 Example of CWMP message interaction



The whole process is as follows:

1. Establish a TCP connection.
2. SSL initialization, and establish a security connection.
3. The CPE sends an Inform request message to initiate a CWMP connection. The Inform message carries the reason for sending this message in the Eventcode field. In this example, the reason is "6 CONNECTION REQUEST", indicating that the ACS requires the CPE to establish a connection.
4. If the CPE passes the authentication of the ACS, the ACS returns an Inform response, and the connection is established.
5. Receiving the Inform response, the CPE sends an empty message, if it has no other requests. The CPE does this in order to comply with the request/reply interaction model of HTTP, in which CWMP messages are conveyed.
6. The ACS queries the value of the ACS URL set on the CPE.
7. The CPE replies to the ACS with the obtained value of the ACS URL.
8. The ACS finds that its local URL value is the same as the value of the ACS URL on the CPE. Therefore, the ACS sends a Set request to the CPE to modify the ACS URL value of the CPE to the URL of the standby ACS.
9. The setting succeeds and the CPE sends a response.
10. The ACS sends an empty message to notify the CPE that it does not request for any other information from the CPE.
11. The CPE closes the connection.

After this, the CPE will initiate a connection to the standby ACS.

CWMP configuration tasks

Configuring the DHCP server

In a CWMP network, the DHCP server is mainly used to notify the ACS location and authentication information to the ACS. DHCP server configuration includes the following tasks:

- Configuring a DHCP address pool for allocating IP addresses to CPEs.
- Configuring the DNS server.
- Configuring the Option 43 field to notify the ACS information to CPEs.

The following describes how to configure the option 43 field:

You can configure ACS parameters for the CPE on the DHCP server through DHCP Option 43. When accessed by the CPE, the DHCP server sends the ACS parameters in DHCP Option 43 to the CPE. If the DHCP server is an HP switch that supports DHCP Option 43, you can configure the ACS parameters at the CLI with the command **option 43 hex 01 length URL username password**, where:

- *length* is a hexadecimal string that indicates the total length of the *URL username password* arguments. No space is allowed between the **01** keyword and the length value.
- *URL* is the ACS address.
- *username* is the ACS username.
- *password* is the ACS password.

When configuring the ACS URL, username and password, follow these guidelines:

- The three arguments take the hexadecimal format and the ACS URL and username must each end with a space (20 in hexadecimal format) for separation.
- The three arguments must be input in 2-digit, 4-digit, 6-digit, or 8-digit segments, each separated by a space.

For example, to set the ACS address to **http://169.254.76.31:7547/acs**, username to **1234**, and password to **5678**, you can configure as follows:

```
<Sysname> system-view
[Sysname] dhcp server ip-pool 0
[Sysname-dhcp-pool-0] option 43 hex 0127 68747470 3A2F2F31 36392E32 35342E37 362E3331
3A373534 372F6163 73203132 33342035 3637 38
```

In the **option 43 hex** command:

- 27 indicates that the length of the subsequent hexadecimal strings is 39 characters.
- 68747470 3A2F2F31 36392E32 35342E37 362E3331 3A373534 372F6163 73 corresponds to the ACS address **http://169.254.76.31/acs**.
- 3132 3334 corresponds to the username **1234**.
- 35 3637 38 corresponds to the password **5678**.
- 20 is the end delimiter.

For more information about DHCP, DHCP Option 43, the **option** command, DHCP address pool configuration, and DNS server configuration, see *Layer 3—IP Services Configuration Guide*.

Configuring the DNS server

On the DNS server, you need to bind the URL address to the IP address of the ACS server to make sure that CPEs can obtain the IP address of the ACS through the DNS function.

Configuring the ACS server

An ACS performs auto-configuration of a CPE through remote management. For the primary configurable parameters, see "[Configuration parameter deployment](#)." For how to configure the ACS server, see the user manual came with your ACS server.

Configuring CPEs

You can set CWMP parameters at the CLI.

The switches operate as CPEs in a CWMP-enabled network, so the following describes only the configuration on CPEs.

Complete these tasks to configure CWMP:

| Task | Remarks |
|---|----------|
| Enabling CWMP | Required |
| Configuring the ACS server: | |
| Configuring the ACS URL | Required |
| Configuring the ACS username and password | Optional |
| Configuring CPE attributes: | |
| Configuring the CPE username and password | Optional |
| Configuring the CWMP connection interface | Optional |
| Configuring the CWMP connection interface | Optional |
| Configuring the maximum number of attempts made to retry a connection | Optional |
| Configuring the close-wait timer of the CPE | Optional |
| Configuring the CPE working mode | Optional |
| Specifying an SSL client policy for HTTPS connection to ACS | Optional |

Enabling CWMP

CWMP configurations can take effect only after you enable CWMP.

To enable CWMP:

| Step | Command | Remarks |
|-----------------------|--------------------|---------|
| 1. Enter system view. | system-view | N/A |
| 2. Enter CWMP view. | cwmp | N/A |

| Step | Command | Remarks |
|-----------------|--------------------|---|
| 3. Enable CWMP. | cwmp enable | Optional.
By default, CWMP is enabled. |

Configuring the ACS attributes

ACS attributes include ACS URL, username and password. When the CPE initiates a connection to the ACS, the ACS URL, username and password are carried in the connection request. After the ACS receives the request, if the parameter values in the request are consistent with those configured locally, the authentication succeeds, and the connection is allowed to be established. If not, the authentication fails, and the connection is not allowed to be established.

Configuring the ACS URL

You can assign only one ACS to a CPE and the ACS URL you configured overwrites the old one, if any. To configure the ACS URL:

| Step | Command | Remarks |
|---------------------------|--------------------------------|---------------------------------------|
| 1. Enter system view. | system-view | N/A |
| 2. Enter CWMP view. | cwmp | N/A |
| 3. Configure the ACS URL. | cwmp acs url <i>url</i> | By default, no ACS URL is configured. |

Configuring the ACS username and password

To pass ACS authentication, make sure that the configured username and password are the same as those configured for the CPE on the ACS.

To configure the ACS username and password:

| Step | Command | Remarks |
|--|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter CWMP view. | cwmp | N/A |
| 3. Configure the ACS username for connection to the ACS. | cwmp acs username <i>username</i> | By default, no ACS username is configured for connection to the ACS.

Optional. |
| 4. Configure the ACS password for connection to the ACS. | cwmp acs password [cipher simple] <i>password</i> | You can specify a username without a password that is used in the authentication. If so, the configuration on the ACS and that on the CPE must be the same.

By default, no ACS password is configured for connection to the ACS. |

Configuring CPE attributes

CPE attributes include CPE username and password, which are used by a CPE to authenticate an ACS. When an ACS initiates a connection to a CPE, the ACS sends a session request carrying the CPE URL, username, and password. When the switch (CPE) receives the request, it compares the CPE URL, username, and password with those configured locally. If they are the same, the ACS passes the authentication of the CPE, and the connection establishment proceeds. Otherwise, the authentication fails, and the connection establishment is terminated.

Configuring the CPE username and password

| Step | Command | Remarks |
|--|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter CWMP view. | cwmp | N/A |
| 3. Configure the CPE username for connection to the CPE. | cwmp cpe username <i>username</i> | By default, no CPE username is configured for connection to the CPE. |
| 4. Configure the CPE password for connection to the CPE. | cwmp cpe password [cipher simple] <i>password</i> | Optional.
You can specify a username without a password that is used in the authentication. If so, the configuration on the ACS and that on the CPE must be the same.
By default, no CPE password is configured for connection to the CPE. |

Configuring the CWMP connection interface

The CWMP connection interface is the interface that the CPE uses to communicate with the ACS. The CPE sends the IP address of this interface in the Inform messages and the ACS replies to this IP address for setting up a CWMP connection.

If the interface that connects the CPE to the ACS is the only Layer 3 interface that has an IP address on the device, you do not need to specify the CWMP connection interface. If multiple Layer 3 interfaces are configured, specify the CWMP connection interface to make sure that the IP address of the interface connects to the ACS is sent to the ACS for setting up CWMP connection.

To configure a CWMP connection interface:

| Step | Command | Remarks |
|--|--|--|
| 1. Enter system view. | system-view | N/A |
| 2. Enter CWMP view. | cwmp | N/A |
| 3. Set the interface that connects the CPE to the ACS. | cwmp cpe connect interface <i>interface-type interface-number</i> | By default, the interface for connecting the CPE to the ACS is not configured. |

Sending Inform messages

You must configure the Inform message sending parameter on the CPE to initiate a connection to the ACS.

To configure the CPE to periodically send Inform messages:

| Step | Command | Remarks |
|--|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter CWMP view. | cwmp | N/A |
| 3. Enable the periodical sending of Inform messages. | cwmp cpe inform interval enable | By default, this function is disabled. |
| 4. Configure the interval between sending the Inform messages. | cwmp cpe inform interval seconds | Optional.
By default, the CPE sends an Inform message every 600 seconds. |

To configure the CPE to send an Inform message at a specific time:

| Step | Command | Remarks |
|--|----------------------------------|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter CWMP view. | cwmp | N/A |
| 3. Configure the CPE to send an Inform message at a specific time. | cwmp cpe inform time time | By default, no time is set. The CPE is not configured to send an Inform message at a specific time. |

Configuring the maximum number of attempts made to retry a connection

If a CPE fails to establish a connection to an ACS, or the connection is interrupted during the session (the CPE does not receive a message indicating the normal close of the session), the CPE can automatically reinitiate a connection to the ACS.

To configure the maximum number of attempts that a CPE can make to retry a connection:

| Step | Command | Remarks |
|--|-------------------------------------|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter CWMP view. | cwmp | N/A |
| 3. Configure the maximum number of attempts that a CPE can make to retry a connection. | cwmp cpe connect retry times | Optional.
By default, a CPE regularly sends connection requests to the ACS until a connection is set up. |

Configuring the close-wait timer of the CPE

The close-wait timeout is used mainly in the following cases:

- During the establishment of a connection: If the CPE sends connection requests to the ACS, but the CPE does not receive a response within the configured close-wait timeout, the CPE will consider the connection failed.
- After a connection is established: If there is no packet interaction between the CPE and ACS within the configured close-wait timeout, the CPE will consider the connection invalid, and disconnect the connection.

To configure the close wait timer of a CPE:

| Step | Command | Remarks |
|---|--------------------------------------|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter CWMP view. | cwmp | N/A |
| 3. Configure the timeout value of the CPE close-wait timer. | cwmp cpe wait timeout seconds | Optional.
The default setting is 30 seconds. |

Configuring the CPE working mode

Configure the device to operate in one of the following CPE modes depending on its position in the network:

- **Gateway mode**—Enables the ACS to manage the device and any CPE attached to the device. Use this mode if the device is the egress to the WAN and has lower-level CPEs.
- **Device mode**—If no CPEs are attached to the device, configure the device to operate in device mode.

Disable CWMP before you change the CPE working mode.

To configure the working mode of the CPE:

| Step | Command | Remarks |
|---|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Enter CWMP view. | cwmp | N/A |
| 3. Configure the working mode of the CPE. | cwmp device-type { device gateway } | By default, the device operates in gateway mode. |

Specifying an SSL client policy for HTTPS connection to ACS

CWMP uses HTTP or HTTPS for data transmission. If the ACS uses HTTPS for secure access, its URL begins with **https://**. You must configure an SSL client policy for the CPE to authenticate the ACS for establishing an HTTPS connection. For more information about configuring SSL client policies, see *Security Configuration Guide*.

To specify an SSL client policy for the CPE to establish an HTTPS connection to the ACS:

| Step | Command | Remarks |
|-----------------------|--------------------|---------|
| 1. Enter system view. | system-view | N/A |
| 2. Enter CWMP view. | cwmp | N/A |

| Step | Command | Remarks |
|----------------------------------|---|--|
| 3. Specify an SSL client policy. | ssl client-policy <i>policy-name</i> | By default, SSL client policy is disabled. |

Displaying and maintaining CWMP

| Task | Command | Remarks |
|-------------------------------------|--|-----------------------|
| Display CWMP configuration. | display cwmp configuration [{ begin exclude include } <i>regular-expression</i>] | Available in any view |
| Display the current status of CWMP. | display cwmp status [{ begin exclude include } <i>regular-expression</i>] | Available in any view |

CWMP configuration example

Configuration guidelines

Before configuring the ACS server, make sure that the HP iMC BIMS software is installed on the server. The BIMS functions and web interface might change along with software updates. If your web interface is different from that in this example, see the user manual came with your server.

Network requirements

A data center has two equipment rooms A and B. Both rooms require a great number of switches. There are ACS, DHCP, and DNS servers on the network. To improve deployment efficiency, use CWMP to deliver different configuration files to the switches in rooms A and B. In this example, each room has three switches.

Figure 63 Network diagram

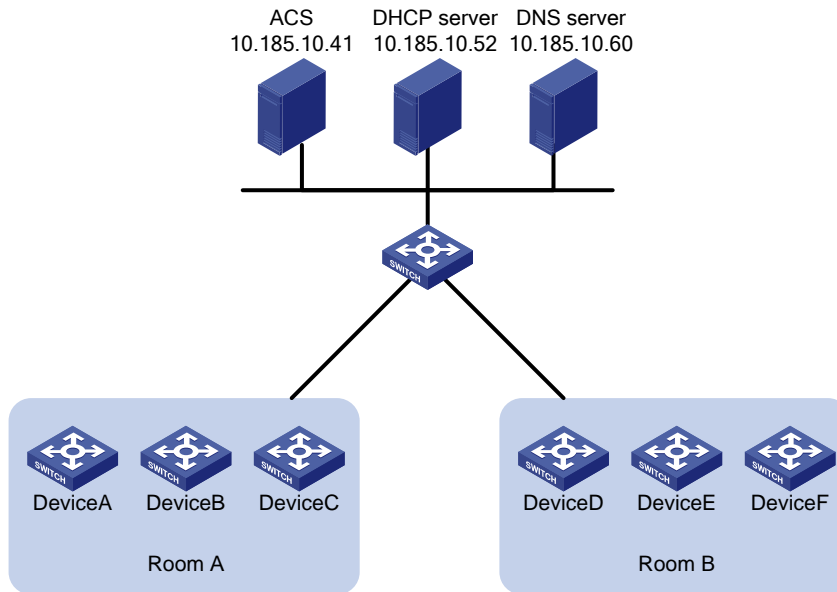


Table 10 Switches deployed in two equipment rooms

| Equipment room | Switch | Serial ID |
|----------------|---------|---------------------|
| A | DeviceA | 210235AOLNH12000008 |
| | DeviceB | 210235AOLNH12000010 |
| | DeviceC | 210235AOLNH12000015 |
| B | DeviceD | 210235AOLNH12000017 |
| | DeviceE | 210235AOLNH12000020 |
| | DeviceF | 210235AOLNH12000022 |

The network administrator has created two configuration files **sys.a.cfg** and **sys_b.cfg** for the switches in the two rooms. The username and password for accessing the ACS server is **vicky** and **12345**. The URL address is <http://acs.database:9090/acs>.

Configuration procedure

1. Configure the ACS server:

ACS server configuration includes the following tasks:

- Setting the username and password for accessing the ACS server.
- Adding information about CPEs and divide CPEs into different groups.
- Binding configuration files to different CPE groups.

Other configurations on the ACS server keep their default value.

Set a username and password on the ACS server.

Click the **System Management** tab, and select **CPE Authentication Users** from the navigation tree to enter the CPE authentication user configuration page.

Figure 64 CPE Authentication User page

System Management >> CPE Authentication User ? Help

Query CPE Authentication Users

Username

CPE Authentication User List

1-2 of 2. Page 1 of 1. Items per Page: 8 15 [50] 100 200

| Username ▾ | Description | Modify | Delete |
|------------|--------------------------------------|--------|--------|
| lif | | | |
| bims | The default CPE authentication user. | | |

Click **Add** to enter the page for adding a CPE authentication user.

Figure 65 Add CPE Authentication User page

System Management >> CPE Authentication User >> Add CPE Authentication User ? Help

Add CPE Authentication User

* Username ?

* Password ?

Description

Set the username, password, and description, and then click **OK**.

Add a device group and a device class. In this example, add **DeviceA** to the **Device_A** class of the **DB_1** group.

Click the **Resource** tab, and select **Group Management > Device Group** from the navigation tree to enter the device group page. Click **Add** to enter the page for adding a device group.

Figure 66 Add Device Group page

Resource >> Device Group >> Add Device Group ? Help

Add Device Group

Basic Info of Device Group

* Group Name ?

Description ?

Operators

| | User Name | Full Name | Role | Manage All Groups | Description |
|-------------------------------------|-----------|-----------|-------|-------------------|--------------------------------|
| <input checked="" type="checkbox"/> | admin | admin | ADMIN | Yes | ADMIN, who has all privileges. |

Set the group name and click **OK**.

Select **Device Class** from the navigation tree. On the device class page, click **Add** to enter the page for adding a device class.

Figure 67 Add Device Class page

Resource >> Device Class >> Add Device Class ? Help

Add Device Class

* Class Name ?

Class Description ?

After setting the class name, click **OK**.

Select **Add Device** from the navigation tree to enter the page for adding a device.

Figure 68 Add Device page

Resource >> Add Device ? Help

Add Device

* Device Name ?

Vendor ?

* OUI ?

* Serial ID ?

Device Class ▼

Device Group ▼

Input the device information and click **OK**.

Figure 69 Adding device succeeded

Resource >> All Devices ? Help

Adding device "DeviceA" succeeded.

Query Device

Device Name Serial ID

Device Class Device Status ▼

Vendor IP Address

Device List

1-50 of 203. Page 1 of 5. Items per Page: 8 15 **50** 100 200

| <input type="checkbox"/> | Status | Device Name | NAT Device | Serial ID | Device Class | Vendor | IP Address | Operation |
|--------------------------|---------|-------------|------------|---------------------|--------------|--------|------------|-----------|
| <input type="checkbox"/> | Unknown | DeviceA | No | 210235AOLNH12000008 | Device_A | HP | | |

Repeat the previous steps to add information about DeviceB and DeviceC to the ACS server, and the adding operation of switches in equipment room A is completed.

Bind different configuration files to different CPE groups to realize auto-deployment.

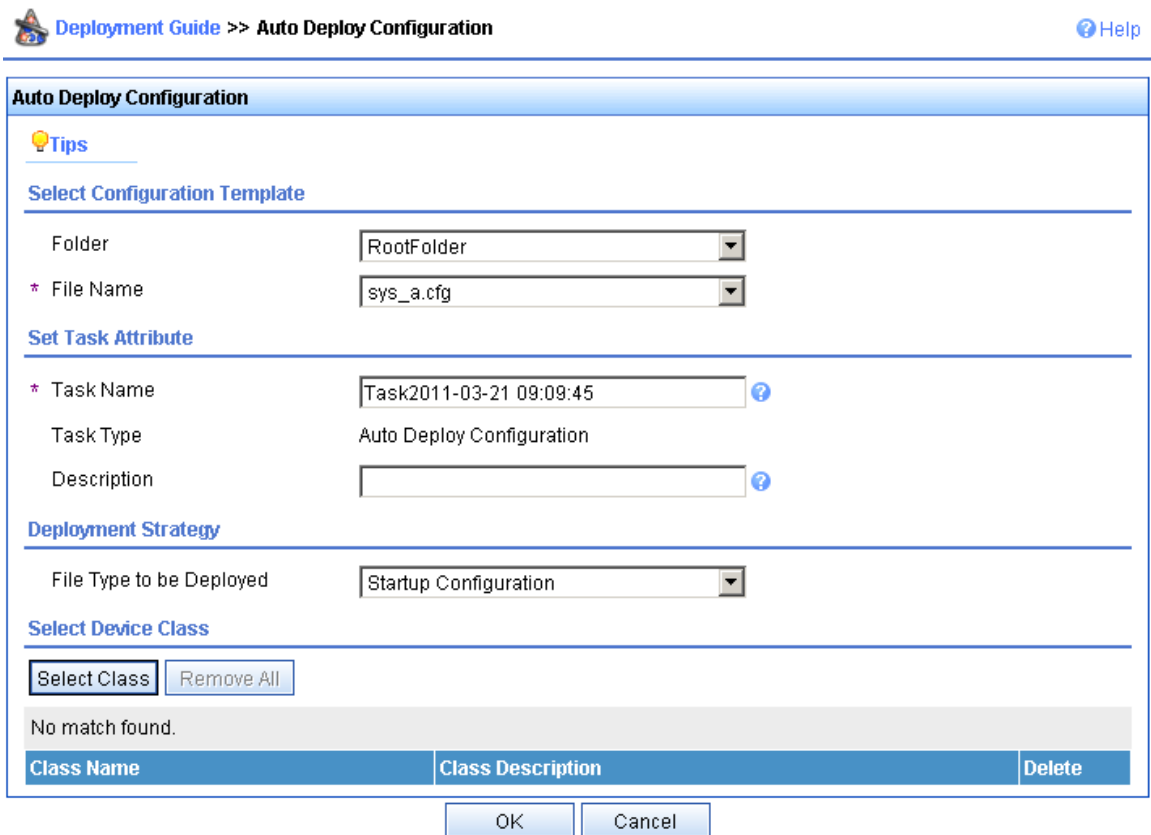
Select **Deployment Guide** from the navigation tree. On the deployment guide page, select **By Device Type** in the **Auto Deploy Configuration** area.

Figure 70 Deployment Guide page



On the **Auto Deploy Configuration** page, select the configuration file to be deployed and set it as the startup configuration as the deployment strategy.

Figure 71 Auto Deploy Configuration page



Click **Select Class** and enter the page for selecting device type.

Figure 72 Selecting a device class

Device Class

Class Name

Class List

1-4 of 4. Page 1 of 1. Items per Page: 8 15 **[50]** 100 200

| <input type="checkbox"/> | Class Name | Class Description |
|-------------------------------------|------------|-------------------|
| <input checked="" type="checkbox"/> | Device_A | |
| <input type="checkbox"/> | MSR50-40 | |
| <input type="checkbox"/> | MSR50-20 | |
| <input type="checkbox"/> | MSR30-40 | |

Select the **Device_A** device class and click **OK**. After that, the auto deploy configuration page is displayed. Click **OK** to complete the task.

Figure 73 Deploying task succeeded

Deployment Task [Help](#)

✔ Creating task "Task2011-03-21 09:09:45" succeeded.

Query Condition

Task Name Task Type

 Task Status Operation Result

Deployment Task List

Run
Suspend
Resume
Delete
Refresh

1-5 of 5. Page 1 of 1. Items per Page: 8 15 **[50]** 100 200

| <input type="checkbox"/> | Task Status--Operation Result | Task Name | Task Type | Creation Time | Creator | Start Time | End Time | Modify | Copy | Delete |
|--------------------------|-------------------------------|-------------------------|---------------------------|---------------------|---------|------------|----------|--------|------|--------|
| <input type="checkbox"/> | Waiting--Unknown | Task2011-03-21 09:09:45 | Auto Deploy Configuration | 2012-05-21 09:10:45 | admin | -- | -- | | | |

Configuration of the switches in room B is the same as that of the switches in room A except that you need to perform the following configuration:

- Create device class Device_B for switches in room B.
- Add switches in room B to the device class Device_B.
- Bind the configuration file corresponding to switches in room B to the device class Device_B.

2. Configure the DHCP server:

In this example, the DHCP server is an HP switch supporting the Option 43 function. If your DHCP server is not an HP switch supporting the Option 43 function, see the user manual came with your server.

Configure a DHCP address pool. Assign IP addresses to CPEs and the DNS server. In this example, the addresses are in the network segment 10.185.10.0/24.

Enable DHCP.

```
<DHCP_server> system-view
[DHCP_server] dhcp enable
```

Enable the DHCP server on VLAN-interface 1.

```
[DHCP_server] interface vlan-interface 1
[DHCP_server-Vlan-interface1] dhcp select server global-pool
[DHCP_server-Vlan-interface1] quit
```

Exclude IP addresses (addresses of the DNS server and ACS server).

```
[DHCP_server] dhcp server forbidden-ip 10.185.10.41
[DHCP_server] dhcp server forbidden-ip 10.185.10.60
```

Configure DHCP address pool 0 (subnet and DNS server address).

```
[DHCP_server] dhcp server ip-pool 0
[DHCP_server-dhcp-pool-0] network 10.185.10.0 mask 255.255.255.0
[DHCP_server-dhcp-pool-0] dns-list 10.185.10.60
```

Configure Option 43 to contain the ACS address, username, and password.

Convert the ACS address, username, and password to ASCII code. The ASCII code of the URL address is 68 74 74 70 3A 2F 2F 61 63 73 2E 64 61 74 61 62 61 73 65 3A 39 30 39 30 2F 61 63 73, that of the username Vicky is 76 69 63 6B 79, and that of the password 12345 is 31 32 33 34 35.

```
[DHCP_server-dhcp-pool-0] option 43 hex 0140 68747470 3A2F2F61 63732E64 61746162
6173653A 39303930 2F616373 20766963 6B792031 32333435
```

3. Configure the DNS server:

Configure the mappings between the domain name and IP address, that is, create the mapping between the addresses <http://acs.database:9090/acs> and <http://10.185.1.41:9090/acs>. For how to create a mapping between addresses, see the user manual came with your DNS server.

4. Connect CPEs to the network:

Connect the CPEs with network cables and power them on, the CPEs can automatically obtain configuration files from the ACS server.

5. Verify the configuration on the ACS server:

Click the **Resource** tab, select **Device Interaction Log** from the navigation tree to enter the page for querying device interaction records. You can view whether the deployment configuration of a switch is completed.

Figure 74 Device Interaction Log page

 [Resource](#) >> **Device Interaction Log** [Help](#)

Query Interaction Log

Device Name Description

Start time   End time  

Interaction Log List

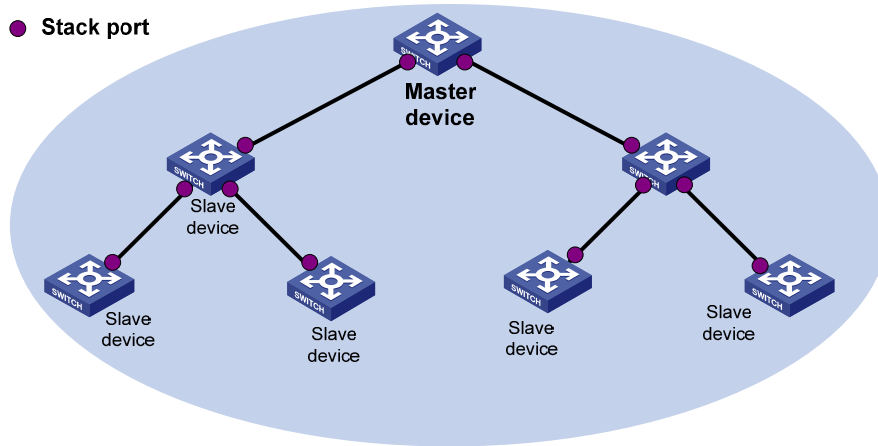
1-15 of 15. Page 1 of 1. Items per Page: [3](#) [15](#) **[50](#)** [100](#) [200](#)

| Device Name | OUI | Serial ID | IP Address | Operation Time ▾ | Description |
|-------------|-----|-----------|------------|------------------|-------------|
|-------------|-----|-----------|------------|------------------|-------------|

Configuring a stack (available only on the HP 5500 SI)

The stack management feature enables you to configure and monitor a group of connected switches by logging in to one switch in the stack, as shown in [Figure 75](#).

Figure 75 Network diagram for stack management



To set up a stack for a group of connected switches, you must log in to one switch to create the stack. This switch is the master switch for the stack, and you configure and monitor all other member switches on the master switch. The ports that connect the stack member switches are called stack ports.

Stack configuration task list

| Task | Remarks |
|---|----------|
| Configuring the master device of a stack | |
| Configuring a private IP address pool for the stack | Required |
| Configuring stack ports | Required |
| Creating a stack | Required |
| Configuring stack member switches | |
| Configuring the stack ports of a member device | Required |
| Logging in to the CLI of a member from the master | Optional |

Configuring the master device of a stack

Perform the tasks in this section to configure the master device. After you complete the stack configuration, the master automatically add member devices to the stack.

Always start configuring the master device with assigning a private IP address pool for the stack. You cannot perform this task after the switch is configured as the master device or a member device.

Configuring a private IP address pool for the stack

Make sure that the number of IP addresses in the address pool is equal to or greater than the number of devices to be added to the stack. If not, some devices cannot automatically join the stack for lack of private IP addresses.

To configure a private IP address pool for the stack:

| Step | Command | Remarks |
|---|---|---|
| 1. Enter system view. | system-view | N/A |
| 2. Configure a private IP address pool for the stack. | stack ip-pool <i>ip-address</i> { <i>mask</i> <i>mask-length</i> } | By default, no IP address pool is configured for a stack. |

Configuring stack ports

Configure the ports that connect the master to member devices as stack ports.

To configure stack ports:

| Step | Command | Remarks |
|--|--|---------------------------------------|
| 1. Enter system view. | system-view | N/A |
| 2. Configure the specified ports as stack ports. | stack stack-port <i>stack-port-num</i>
port <i>interface-list</i> | By default, no ports are stack ports. |

Creating a stack

After you execute the **stack role master** command, the device becomes the master device of a stack and automatically adds the devices connected to its stack ports to the stack.

To create a stack:

| Step | Command |
|-----------------------|--------------------------|
| 1. Enter system view. | system-view |
| 2. Create a stack. | stack role master |

After you configure a device as the master device of a stack, the prompt changes to <stack_0.Sysname>, where Sysname is the system name of the device.

Configuring the stack ports of a member device

To add a device to a stack, you must configure the ports that connect the device to other stack members as stack ports.

To configure stack ports:

| Step | Command | Remarks |
|--|--|---|
| 1. Enter system view. | system-view | N/A |
| 2. Configure the specified ports as stack ports. | stack stack-port stack-port-num
port interface-list | By default, a port is not a stack port. |

After a device joins a stack and becomes a member device of the stack, the prompt changes to <stack_n.Sysname>, where n is the stack number assigned by the master device, and Sysname is the system name of the device.

Logging in to the CLI of a member from the master

Perform this task on the master device in user view.

| Task | Command |
|--|----------------------------------|
| Log in to the CLI of a member device from the master device. | stack switch-to member-id |

The **stack switch-to** command does not change the user privilege level. To return to the master device, use the **quit** command.

Displaying and maintaining stack configuration

| Task | Command | Remarks |
|---|---|-----------------------|
| Display the stack configuration of stack members. | display stack [members] [
{ begin exclude include }
regular-expression] | Available in any view |

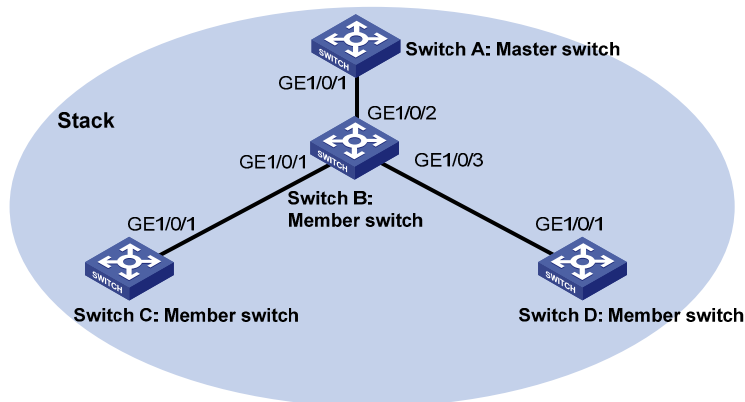
Stack configuration example

Network requirements

Switch A, Switch B, Switch C, and Switch D are connected with one another.

Create a stack, where Switch A is the master device, Switch B, Switch C, and Switch D are member devices. An administrator can log in to Switch B, Switch C and Switch D through Switch A to perform remote configurations.

Figure 76 Network diagram



Configuration procedure

1. Configure the master device:

Configure a private IP address pool for the stack on Switch A.

```
<SwitchA> system-view
```

```
[SwitchA] stack ip-pool 192.168.1.1 24
```

Configure port GigabitEthernet 1/0/1 as a stack port on Switch A.

```
[SwitchA] stack stack-port 1 port gigabitethernet 1/0/1
```

Configure switch A as the master device.

```
[SwitchA] stack role master
```

2. Configure the member devices:

On Switch B, configure local ports GigabitEthernet 1/0/2, GigabitEthernet 1/0/1, and GigabitEthernet 1/0/3 as stack ports.

```
<SwitchB> system-view
```

```
[SwitchB] stack stack-port 3 port gigabitethernet 1/0/1 gigabitethernet 1/0/2  
gigabitethernet 1/0/3
```

On Switch C, configure local port GigabitEthernet 1/0/1 as a stack port.

```
<SwitchC> system-view
```

```
[SwitchC] stack stack-port 1 port gigabitethernet 1/0/1
```

On Switch D, configure local port GigabitEthernet 1/0/1 as a stack port.

```
<SwitchD> system-view
```

```
[SwitchD] stack stack-port 1 port gigabitethernet 1/0/1
```

3. Verify the configuration:

Display information about stack members on Switch A.

```
<stack_0.SwitchA> display stack members
```

```
Number      : 0
```

```
Role        : Master
```

```
Sysname     : stack_0. SwitchA
```

```
Switch type: HP 5500-24G SI Switch with 2 Interface Slots
```

```
MAC address: 000f-e200-1000
```

```
Number      : 1
```

```
Role        : Slave
```

```
Sysname     : stack_1. SwitchB
```

Device type: HP 5500-24G SI Switch with 2 Interface Slots
MAC address: 000f-e200-1001

Number : 2
Role : Slave
Sysname : stack_2. DeviceC
Device type: HP 5500-24G SI Switch with 2 Interface Slots
MAC address: 000f-e200-1002

Number : 3
Role : Slave
Sysname : stack_3. DeviceD
Device type: HP 5500-24G SI Switch with 2 Interface Slots
MAC address: 000f-e200-1003

Index

[A](#) [C](#) [D](#) [E](#) [H](#) [I](#) [L](#) [N](#) [O](#) [P](#) [S](#) [T](#) [U](#)

A

- Adding a candidate switch to a cluster, [178](#)
- Alarm group configuration example, [78](#)

C

- Cluster management configuration example, [182](#)
- Cluster management configuration task list, [169](#)
- Configuring a schedule for an NQA test group, [124](#)
- Configuring access between the management switch and its member switches, [177](#)
- Configuring access-control rights, [18](#)
- Configuring advanced cluster management functions, [178](#)
- Configuring an NQA test group, [108](#)
- Configuring CPE attributes, [194](#)
- Configuring NTP authentication, [19](#)
- Configuring NTP operation modes, [13](#)
- Configuring optional parameters, [17](#)
- Configuring optional parameters for an NQA test group, [123](#)
- Configuring PoE interface through PoE profile, [159](#)
- Configuring PoE power management, [158](#)
- Configuring sFlow, [147](#)
- Configuring the ACS attributes, [193](#)
- Configuring the collaboration function, [119](#)
- Configuring the history records saving function, [122](#)
- Configuring the management switch, [170](#)
- Configuring the master device of a stack, [206](#)
- Configuring the member switches, [177](#)
- Configuring the NQA server, [107](#)
- Configuring the NQA statistics collection function, [122](#)
- Configuring the PoE monitoring function, [159](#)
- Configuring the PoE power, [157](#)
- Configuring the RMON alarm function, [74](#)
- Configuring the RMON statistics function, [73](#)
- Configuring the stack ports of a member device, [207](#)
- Configuring threshold monitoring, [120](#)
- Creating an NQA test group, [108](#)
- CWMP configuration example, [197](#)

- CWMP configuration tasks, [191](#)
- CWMP overview, [186](#)

D

- Detecting PDs, [157](#)
- Disabling an interface from generating link up/down logging information, [49](#)
- Displaying and maintaining cluster management, [181](#)
- Displaying and maintaining CWMP, [197](#)
- Displaying and maintaining information center, [50](#)
- Displaying and maintaining IPC, [153](#)
- Displaying and maintaining NQA, [125](#)
- Displaying and maintaining NTP, [21](#)
- Displaying and maintaining PoE, [161](#)
- Displaying and maintaining port mirroring, [92](#)
- Displaying and maintaining RMON, [75](#)
- Displaying and maintaining sFlow, [148](#)
- Displaying and maintaining stack configuration, [208](#)

E

- Enabling CWMP, [192](#)
- Enabling IPC performance statistics, [152](#)
- Enabling PoE, [156](#)
- Enabling the NQA client, [107](#)
- Ethernet statistics group configuration example, [76](#)

H

- Hardware compatibility, [186](#)
- History group configuration example, [76](#)

I

- Information center configuration examples, [50](#)
- Information center configuration task list, [39](#)
- Introduction to port mirroring, [81](#)

L

- Logging in to the CLI of a member from the master, [208](#)

N

- NQA configuration examples, [126](#)
- NQA configuration task list, [106](#)

NTP configuration examples, [21](#)

NTP configuration task list, [13](#)

O

Outputting system information to a log host, [42](#)

Outputting system information to the console, [40](#)

Outputting system information to the log buffer, [43](#)

Outputting system information to the monitor terminal, [41](#)

Outputting system information to the SNMP module, [44](#)

Outputting system information to the trap buffer, [43](#)

Outputting system information to the Web interface, [45](#)

Overview, [71](#)

Overview, [154](#)

Overview, [8](#)

Overview, [103](#)

Overview, [151](#)

Overview, [33](#)

Overview, [165](#)

P

Ping, [1](#)

Ping and tracer example, [7](#)

PoE configuration example, [162](#)

PoE configuration task list, [155](#)

Port mirroring configuration examples, [92](#)

S

Saving security logs into the security log file, [45](#)

sFlow configuration example, [149](#)

sFlow overview, [146](#)

Stack configuration example, [208](#)

Stack configuration task list, [206](#)

System debugging, [5](#)

T

Tracert, [3](#)

Troubleshooting PoE, [163](#)

Troubleshooting sFlow configuration, [150](#)

U

Upgrading PSE processing software in service, [161](#)